VOLUME 64
NUMBER 8 / NUMÉRO 8
AUGUST / AOÛT
2007

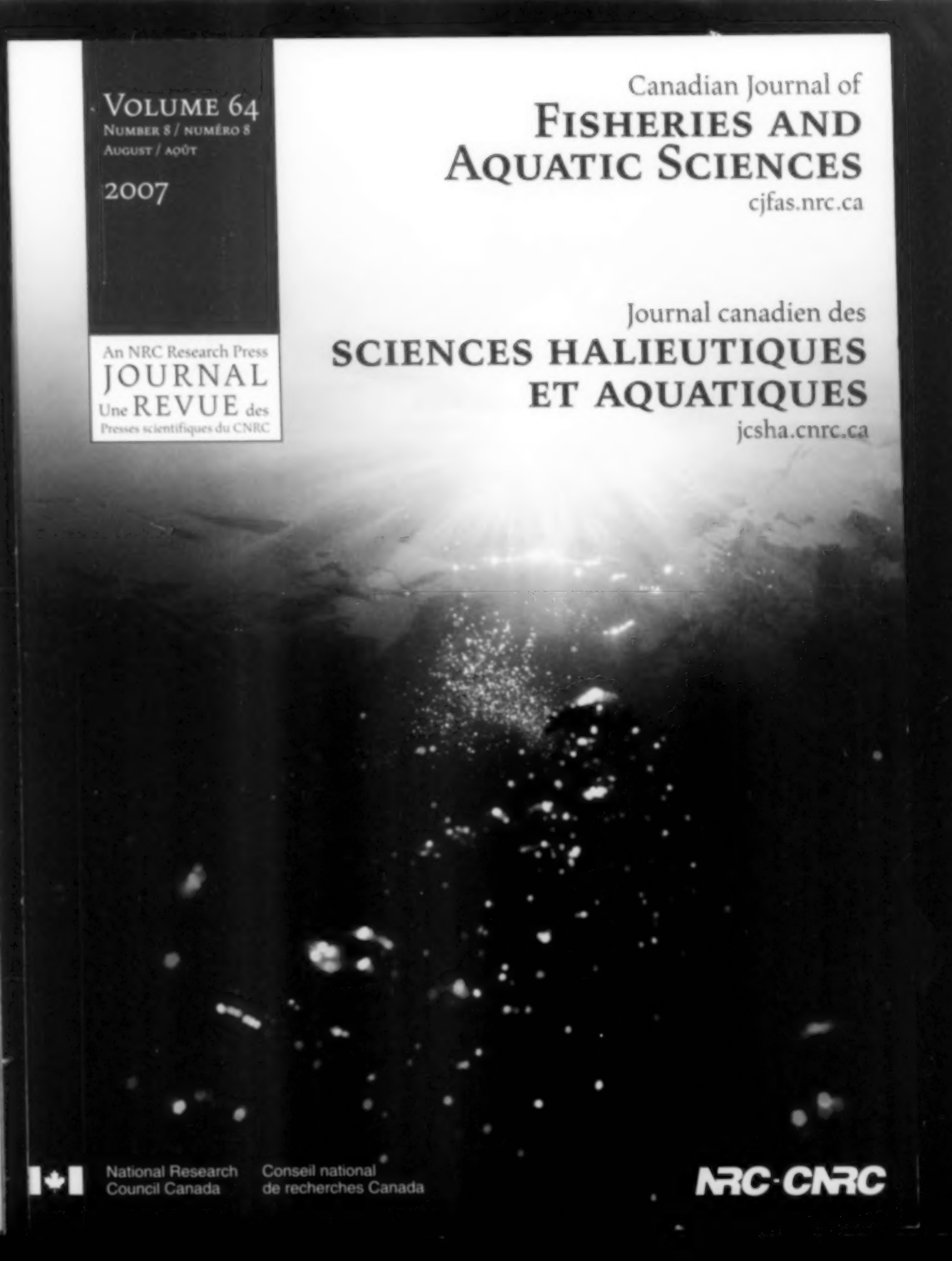
Canadian Journal of FISHERIES AND AQUATIC SCIENCES

cjfas.nrc.ca

Journal canadien des SCIENCES HALIEUTIQUES ET AQUATIQUES

jcsha.cnrc.ca

An NRC Research Press
JOURNAL
Une REVUE des
Presses scientifiques du CNRC



National Research Council Canada
Conseil national de recherches Canada
NRC·CNRC

ISSN 0706-652X / CJFSDX 64(8) 1043–1155 (2007)

The *Canadian Journal of Fisheries and Aquatic Sciences*, previously published as *Contributions to Canadian Biology* 1901–1925, *Contributions to Canadian Biology and Fisheries* 1926–1934, *Journal of the Biological Board of Canada* 1934–1937, and *Journal of the Fisheries Research Board of Canada* 1938–1979, is issued monthly to form annual volumes of 12 numbers plus supplements.

Editorial policy. The Journal publishes original research articles, perspectives (syntheses, critiques, reevaluations), discussions (comment and reply), and rapid communications. Manuscripts may concern cells, organisms, populations, communities, ecosystems, or processes that affect aquatic systems. They may cover a range of disciplines including biology and ecology of marine and freshwater organisms, limnology, oceanography, physiology, toxicology, genetics, economics, disease, and management. Preference will be given to manuscripts that emphasize understanding of observed phenomena and interpretation of experimental results. The Journal encourages papers that lead from clearly stated purpose or rationale and from testable hypotheses, concepts, or questions to identifiable conclusions or syntheses. Such papers may amplify, modify, question, or redirect accumulated knowledge embodied in contemporary perceptions of a particular state of aquatic sciences. The Journal discourages papers that are essentially descriptive, except in emerging disciplines; have only site-specific or local applications (a certain year, place, taxon, chemical compound, etc.); only confirm previously established principles; or apply standard techniques without breaking new methodological ground.

Manuscripts for publication should be submitted to the Editorial Office via NRC Research Press' online submission and peer review system (**OSPREY**) at http://pubs.nrc-cnrc.gc.ca/cjfas/osprey. Alternatively, manuscripts may be submitted to the Editor, *Canadian Journal of Fisheries and Aquatic Sciences*, Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON M5S 3G5, Canada (416-946-0609; fax: 416-946-0509).

Submissions in either English or French are acceptable. The information must be original, that is, not copyrighted, published, or submitted elsewhere except in abstract form or unless by written consent of the Editor.

Correspondence concerning accepted manuscripts and corrected proofs should be sent to NRC Research Press, National Research Council of Canada, Ottawa, ON K1A 0R6, Canada (fax: 613-952-7656; e-mail: pubs@nrc-cnrc.gc.ca).

Instructions to Authors are published once a year in the first issue of each volume. Copies are available from NRC Research Press or at cjfas.nrc.ca. The **Publication Policy** of NRC Research Press is available at pubs.nrc-cnrc.gc.ca.

Microfilm. Issues from 1934 through the current volume can be purchased on 16- or 35-mm microfilm. Photocopies of individual articles or issues can be purchased from University Microfilms International, 330 North Zeeb Road, Ann Arbor, MI 48106, USA.

The *Canadian Journal of Fisheries and Aquatic Sciences*, ISSN 0706-652X, is published monthly, 12 times per year, by NRC Research Press, c/o USACAN Media Dist. Srv. Corp. at 26 Power Dam Way, Suite S1-S3, Plattsburgh, NY 12901, USA. Periodicals postage paid at Plattsburgh, NY. POSTMASTER: Send address changes to *Canadian Journal of Fisheries and Aquatic Sciences*, c/o NRC, P.O. Box 2888, Plattsburgh, NY 12901, USA.

Publié sans interruption depuis 1901, le *Journal canadien des sciences halieutiques et aquatiques* a paru sous plusieurs titres : *Contributions to Canadian Biology* 1901–1925, *Contributions to Canadian Biology and Fisheries* 1926–1934, *Journal of the Biological Board of Canada* 1934–1937, et *Journal of the Fisheries Research Board of Canada* 1938–1979. Le Journal paraît tous les mois, formant ainsi un volume de 12 numéros, et offre en plus jusqu'à deux suppléments par année.

Politique éditoriale. On publie, dans le Journal, des articles rendant compte de recherche originale, des perspectives (synthèses, critiques et réévaluations), des discussions (commentaires et réponses) et des communications rapides. Les documents peuvent porter sur les cellules, organismes, populations, communautés, écosystèmes ou processus qui affectent les systèmes aquatiques. Ils peuvent relever de divers domaines y compris la biologie et l'écologie des organismes marins et d'eau douce, la limnologie, l'océanographie, la physiologie, la toxicologie, la génétique, l'économie, la pathologie et la gestion. On accorde la préséance aux articles qui contribuent à la compréhension de phénomènes observés et à l'interprétation de résultats expérimentaux. On invite les auteurs à soumettre des articles ayant des objectifs précis ou relevant d'hypothèses, de concepts ou de questions vérifiables et se terminant par de solides conclusions ou d'utiles synthèses. Ces articles enrichissent, changent, mettent en question ou réorientent les connaissances acquises sur divers aspects des sciences aquatiques. On décourage toutefois la soumission d'articles simplement descriptifs (à moins qu'il ne s'agisse de nouvelles disciplines), d'application très limitée (applicables à une seule année, une seule région ou un seul taxon, portant sur un unique composé chimique, etc.); d'articles qui ne font que reprendre les principes énoncés dans des articles précédents; et d'articles rendant compte de l'application de normes sans gain de méthodologie.

Les auteurs sont invités à soumettre leur **manuscrit** au Bureau de la rédaction par le truchement du Système de soumission et d'examen en direct des articles (**OSPREY**) des Presses scientifique du CNRC à http://pubs.nrc-cnrc.gc.ca/jcsha/osprey. Sinon, ils peuvent transmettre leur manuscrit par la poste au directeur scientifique, *Journal canadien des sciences halieutiques et aquatiques*, Département d'écologie et de biologie évolutive, Université de Toronto, Toronto, ON M5S 3G5, Canada (416-946-0609; télécopieur : 416-946-0509).

Les contributions peuvent être en anglais ou en français. Elles doivent être originales, c.-à.-d. qu'elles ne doivent pas avoir fait l'objet d'un copyright, avoir été publiées ou soumises ailleurs, sauf sous forme abrégée ou avec le consentement du rédacteur. Le CNRC ne se tient en aucune façon responsable des opinions exprimées par un collaborateur ou des déclarations qui pourrait faire ce dernier.

La correspondance ayant trait aux manuscrits acceptés et les **épreuves corrigés** doivent être envoyées aux Presses scientifiques du CNRC, Conseil national de recherches du Canada, Ottawa, ON K1A 0R6, Canada (télécopieur : 613-952-7656; courriel : pubs@nrc-cnrc.gc.ca).

Les Recommandations aux auteurs sont publiées une fois par année dans le premier numéro de chaque volume. Des copies sont disponibles auprès du Bureau de la rédaction aux Presses scientifiques du CNRC ou au jcsha.cnrc.ca. **La Politique de publication** est disponible auprès des Presses scientifiques du CNRC ou au pubs.nrc-cnrc.gc.ca.

Microfilms. On peut acheter sur microfilms de 16 ou 35 mm tous les numéros depuis 1934. Des photocopies d'articles ou de numéros individuels sont aussi en vente auprès de University Microfilms International, 330 North Zeeb Road, Ann Arbor, MI 48106, É.-U.

This Journal is indexed or abstracted in / Cette revue est signalée ou résumée dans

ABSEARCH; Applied Science and Technology Index and Abstracts; Aquatic Sciences and Fisheries Abstracts; Biological Abstracts; Biological and Agricultural Abstracts; CAB Abstracts; Chemical Abstracts; Current Contents; Elsevier BIOBASE; Environmental Abstracts; ETDE Energy Database; General Science Index and Abstracts; GEOBASE; GeoRef; Life Sciences Collection; Meteorological and Geoastrophysical Abstracts; Oceanic Abstracts; PASCAL Database; Science Citation Index.

Volume 64, Number 8, August 2007

Canadian Journal of Fisheries and Aquatic Sciences

Published by the National Research Council of Canada

Volume 64, numéro 8, août 2007

Journal canadien des sciences halieutiques et aquatiques

Publié par le Conseil national de recherches du Canada

Editor / Directeur scientifique Don Jackson

Vice-editor / Vice-directeur scientifique

R.D. Vinebrooke, *University of Alberta*

Assistant to the Editors / Adjointe aux directeurs scientifiques **Holly Foster**

Associate Editors / Directeurs scientifiques associés

D. Boisclair, *Université de Montréal*
U. Borgmann, *Environment Canada, Burlington*
M.J. Bradford, *DFO/MPO and Simon Fraser University*
S.E. Campana, *DFO/MPO, Dartmouth*
R. Carignan, *Université de Montréal*
Y. Chen, *University of Maine, Orono, Maine*
K.T. Frank, *DFO/MPO, Dartmouth*
J.W.A. Grant, *Concordia University*
M.M. Hansen, *Danish Institute for Fisheries Research, Silkeborg, Denmark*
R. Hilborn, *University of Washington*
D.F. Houlihan, *University of Aberdeen*
F.A. Huntingford, *University of Glasgow*
J.M. Jech, *NFSC/NOAA, Woods Hole, Mass.*
B. Jonsson, *Norwegian Institute for Nature Research, Oslo, Norway*
K.A. Kidd, *University of New Brunswick*
D.L. MacLatchy, *University of New Brunswick, Saint John*
P. Magnan, *Université du Québec à Trois Rivières*
C.T. Marshall, *University of Aberdeen*
R.B. Millar, *University of Auckland, New Zealand*
W.P. Patterson, *University of Saskatchewan*
Y.T. Prairie, *Université du Québec à Montréal*
T.J. Quinn II, *University of Alaska–Fairbanks, Juneau, Alaska*
T.P. Quinn, *University of Washington*
C.W. Ramcharan, *Laurentian University*
J.S. Richardson, *The University of British Columbia*
K. Rose, *Louisiana State University, Baton Rouge, La.*
B. Sainte-Marie, *DFO/MPO, Mont-Joli*
R.E.H. Smith, *University of Waterloo*
S.J. Smith, *DFO/MPO, Dartmouth*
E.B. Taylor, *The University of British Columbia*
W.F. Vincent, *Université Laval*
C.J. Walters, *The University of British Columbia*
G. Weyhenmeyer, *Swedish University of Agricultural Sciences, Uppsala, Sweden*
B. de Young, *Memorial University of Newfoundland*

NRC Research Press / Presses scientifiques du CNRC

Fax / Télécopieur : 613-952-7656; e-mail / courriel : pubs@nrc-cnrc.gc.ca; URL / URL : pubs.nrc-cnrc.gc.ca

Editor-in-Chief / Directeur scientifique en chef **Bruce P. Dancik**
Assistant Editor-in-Chief / Adjoint au directeur scientifique en chef **Donald S. Mavinic**
Director / Directeur **Cameron D. Macdonald**

Acting Business Manager / Gestionnaire des affaires commerciales par intérim **Mike Boroczki**
Advertising and reprints / Annonces publicitaires et tirés à part **Roxanne Landriault** (613-990-2254)
Subscriptions / Abonnements **Eileen Evans-Nantais** (613-990-7873), **André Séguin** (613-993-9084)

Manager, Journals Program / Gestionnaire du Programme des revues **Carolyn Brown**
Managing Editor / Rédactrice administrative **Cecily M. Pearson** (613-993-9099)
Publication Officers / Agents de publication **Jennifer McColl, Thad Costigan**
Editorial Publication Assistants / Adjoints à la publication **Diana Leck, Luc Julien**

Production Manager / Gestionnaire de la production **Brad Stoppa**
Production Application Support Technician / Technicienne au soutien de la production **Judy Hum**
Desktop Publishing Technician / Technicienne en éditique **Lisette Séguin**
Graphics Technician / Technicien en graphisme **Dario Ferrera**
Web Technician / Technicien Web **Dennis Cafferty**

The National Research Council of Canada accepts no responsibility for statements or opinions expressed by a contributor. Acceptance of an advertisement, announcement, or other material does not imply endorsement by either the Journal's Senior Editor or the National Research Council.

© National Research Council of Canada 2007. All rights reserved.

No part of this publication may be reproduced, stored, or transmitted in any form or by any means, without the written permission of the publisher, except as stated below.

The National Research Council of Canada grants **permission*** to individuals to make single copies of articles for personal use. Any person may reproduce short excerpts from articles in the journal for any purpose that respects the moral rights of authors, provided that the source is fully acknowledged. As a courtesy, the consent of authors of such material should be obtained directly from the author.

Authorization to reproduce items for other than personal use, as stated above, is granted by the National Research Council of Canada provided that the copyright fee of $7.00 per copy is paid directly to Access Copyright, The Canadian Copyright Licensing Agency, One Yonge Street, Suite 800, Toronto, ON M5E 1E5, Canada.

*The above permission does not extend to other kinds of copying, such as copying for general distribution, for advertising, or promotional purposes, for creating new collective works, or for resale. For such copying, arrangements must be made with the publisher, National Research Council of Canada, Ottawa, ON K1A 0R6, Canada.

Subscription information

The *Canadian Journal of Fisheries and Aquatic Sciences* (ISSN 0706-652X) is published monthly by the National Research Council of Canada. The 2007 annual subscription rates (second class postage) for Canada are $1051CAN institutional and $325CAN personal, for US & foreign $1051US institutional and $325US personal. For *air* mail add $150US to foreign *rate*. Canadian customers please include 6% GST (business no. 12149 1807 RT 0275). Subscriptions are entered for the calendar year. Remittances should be made payable to the Receiver General for Canada, credit NRC. Price for regular single issues: institutional, $110.00; personal, $35.00. Supplements will be priced separately. Special rates apply to members of Canadian scientific and engineering societies if they subscribe through their respective societies.

There are no refunds on cancelled subscriptions once delivery has begun. Claims for missing issues should be made within four months (North America) or six months (outside North America) following the date of publication. The publisher expects to supply missing issues only when losses have occurred in transit and when the reserve stock will permit. For change of address and subscription enquiries: Subscription Office, NRC Research Press, National Research Council of Canada, Ottawa, ON K1A 0R6, Canada; tel.: 613-993-9084; fax: 613-952-7656; e-mail: pubs@nrc-cnrc.gc.ca.

Microform backfiles are available from UMI, P.O. Box 1346, Ann Arbor, MI 48106-1346, USA, and Micromedia Ltd., 20 Victoria St., Toronto, ON M5C 2N8, Canada.

All issues of the *Canadian Journal of Fisheries and Aquatic Sciences* since 1996 are available on the Web. Pricing information can be found at pubs.nrc-cnrc.gc.ca.

Le Conseil national de recherches du Canada ne se tient en aucune façon responsable des opinions exprimées par un collaborateur ou des déclarations que pourrait faire ce dernier. La publication d'annonces publicitaires, d'avis et autres n'implique pas l'endossement de leur contenu par la directrice principale de la Revue ou le CNRC.

© Conseil national de recherches du Canada 2007. Tous droits réservés.

Il est interdit de reproduire, de sauvegarder ou de transmettre quelque partie de la présente publication dans quelque format ou de quelque façon sans le consentement écrit de l'éditeur, à l'exception des conditions suivantes.

Le Conseil national de recherches du Canada accorde aux particuliers le droit de faire des copies individuelles pour leur usage personnel. Toute personne peut reproduire de courts extraits des articles de la Revue pour autant que les droits moraux de l'auteur soient respectés et que la mention de provenance soit clairement indiquée. Il est de convenance de demander le consentement de l'auteur avant de faire mention de son travail.

On obtiendra l'autorisation de reproduire des articles à d'autres fins que l'usage personnel (tel qu'on l'a décrit ci-dessus) auprès du Conseil national de recherches du Canada pour autant que des redevances de droits d'auteurs de 7,00 $/copie soient versées au Access Copyright, The Canadian Copyright Licensing Agency, 1, rue Yonge, bureau 800, Toronto, ON M5E 1E5, Canada.

*La permission ci-dessus ne s'applique pas à la reproduction à grande échelle, telle que la reproduction aux fins de distribution générale, publicitaire ou promotionnelle, de compilation de nouveaux ouvrages collectifs ou de revente. Dans de tels cas, il faudra ratifier une entente particulière avec l'éditeur, le Conseil national de recherches du Canada, Ottawa, ON K1A 0R6, Canada.

Renseignements sur l'abonnement

Journal canadien des sciences halieutiques et aquatiques (ISSN 0706-652X) est une publication mensuelle du Conseil national de recherches du Canada. Le prix de l'abonnement annuel en 2007 au Canada (courrier de deuxième classe) est de 1051 $CAN pour les organismes et de 325 $CAN pour les particuliers. À l'étranger (y compris les É.-U.), le prix de l'abonnement est de 1051 $US pour les organismes et de 325 $US pour les particuliers. L'expédition par avion exige des frais de 150 $US en sus du prix d'abonnement à l'étranger. Les clients du Canada doivent ajouter la TPS de 6 % (numéro d'entreprise 12149 1807 RT 0275). Les abonnements sont pour l'année civile. Tout paiement doit être adressé au Receveur général du Canada, crédit du CNRC. Le prix par numéro est de 110 $ pour les organismes et de 35 $ pour les particuliers. Le prix des hors-série varie. Les membres des sociétés canadiennes de sciences et de génie bénéficient de taux préférentiels en s'abonnant via leurs sociétés respectives.

Aucun remboursement ne sera accordé ni aucune annulation acceptée une fois que les livraisons auront commencé. Les réclamations de numéros manquants devront être soumises dans les quatre mois (Amérique du Nord) ou dans les six mois (à l'extérieur de l'Amérique du Nord) suivant la date de publication. L'éditeur ne fournira de numéros manquants que si le document a été perdu pendant l'expédition et que si le stock le permet. Adressez tout changement d'adresse ou toute question quant à l'abonnement au : Bureau des abonnements, Presses scientifiques du CNRC, Conseil national de recherches Canada, Ottawa, ON K1A 0R6, Canada; téléphone : 613-993-9084; télécopieur : 613-952-7656; courriel : pubs@nrc-cnrc.gc.ca.

On peut obtenir les fichiers rétrospectifs sur **microforme** auprès de UMI, P.O. Box 1346, Ann Arbor, MI 48106-1346, É.-U., et auprès de Micromedia Ltée, 20, rue Victoria, Toronto, ON M5C 2N8, Canada.

À partir de 1996, tous les numéros de *Journal canadien des sciences halieutiques et aquatiques* sont accessibles sur le Web. On trouvera des renseignements sur les tarifs à pubs.nrc-cnrc.gc.ca.

Assessing pH changes since pre-industrial times in 51 low-alkalinity lakes in Nova Scotia, Canada

Brian K. Ginn, Brian F. Cumming, and John P. Smol

Abstract: Diatom-based paleolimnological techniques were used to reconstruct lake acidification trends in 51 low-alkalinity Nova Scotia lakes that spanned gradients of dissolved organic carbon (DOC) concentrations and sulphate deposition. Pre-industrial, diatom-inferred pH values of these lakes were <6.8, with 31 lakes having pre-industrial pH < 6.0 and two lakes having pH < 5.5. Lakes in Kejimkujik National Park documented the greatest pH decline (–0.4 pH unit (±0.2)) since the 19th century, whereas those in northern parts of the province (e.g., Cape Breton Highlands National Park) experienced little or no acidification, with a net mean pH decline = –0.1 pH unit (±0.2). While the sulphate deposition and diatom-inferred pH changes have not been as great as those observed in other acidified areas of northeastern North America (e.g., Adirondack region of New York or New England), Nova Scotia lakes have experienced biological changes toward more acidophilous diatom assemblages, especially in lakes with low pre-industrial pH values (currently with high DOC concentrations) located in Kejimkujik National Park, which receives the highest loading of sulphate deposition in Nova Scotia. However, the generally low pre-industrial pH values inferred for most of the study lakes suggest that many of these lakes were somewhat naturally acidic, but acidified further as a result of atmospheric deposition.

Résumé : Des techniques paléolimnologiques basées sur les diatomées nous ont servi à reconstituer les tendances d'acidification lacustre dans 51 lacs de faible alcalinité en Nouvelle-Écosse qui couvrent un gradient de concentrations de carbone organique dissous (DOC) et de précipitations de sulfates. Les valeurs de pH avant l'ère industrielle déduite d'après les diatomées dans ces lacs sont <6,8, avec 31 lacs ayant une valeur de pH < 6,0 et deux un pH < 5,5. Les lacs du parc national de Kejimkujik ont connu le déclin de pH le plus important (–0,4 ± 0,2 unité pH) depuis le 19e siècle, alors que ceux du nord de la province (par ex. du parc national du Cap Breton) n'ont subi que peu ou pas d'acidification, avec une diminution moyenne nette du pH de 0,1 unité pH (± 0,2). Bien que les précipitations de sulfates et les changements de pH d'après l'analyse des diatomées n'aient pas été aussi importants que dans d'autres régions acidifiées du nord-est de l'Amérique du Nord (par ex., la région des Adirondacks de l'état de New York ou la Nouvelle-Angleterre), les lacs de Nouvelle-Écosse ont connu des changements biologiques qui ont mené à des peuplements plus acidophiles de diatomées; ceci est particulièrement vrai des lacs à pH bas avant l'ère industrielle (présentement avec de fortes concentrations de DOC) et situés dans le parc national de Kejimkujik, la région qui reçoit la charge la plus importante de précipitations de sulfates dans la Nouvelle-Écosse. Cependant, les valeurs estimées généralement faibles de pH durant l'ère préindustrielle dans la plupart des lacs étudiés indiquent que ces lacs sont un peu acides naturellement et qu'ils se sont acidifiés davantage à cause des précipitations atmosphériques.

[Traduit par la Rédaction]

Introduction

Acidification of surface waters is one of the foremost environmental problems currently affecting freshwater ecosystems, both in eastern North America and around the world. One area that is believed to be greatly affected by acidic deposition is Nova Scotia, Canada (Fig. 1), an area that has been reported to have some of the lowest surface-water pH values in North America (Kerekes et al. 1982; Morrison 2004). In addition to lakes with low pH values, the average pH of precipitation has decreased by 1.0–1.5 pH units between 1950 and the mid-1980s (Underwood et al. 1982; Lusis and Philips 1990). Most of the mainland portion of Nova Scotia is underlain by nonweathering granite and metamorphic bedrock, which provides little in the way of buffering capacity to these systems (Wiltshire and Machell 1981; Underwood et al. 1982). Probably the most dramatic evidence pointing to impacts of acidic deposition has been the extirpation of Atlantic salmon (*Salmo salar*) from many historic nursery rivers, which has been linked to episodic pulses of very low-pH water during crucial stages of salmonid development (Watt 1987; Watt et al. 2000; Clair et al. 2001).

The first study in Nova Scotia to suggest acidic deposition as an environmental stressor was done by Watt et al. (1979), who surveyed 21 lakes near Halifax and observed a pH de-

Received 21 June 2006. Accepted 12 April 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 1 August 2007. J19379

B.K. Ginn,[1] B.F. Cumming, and J.P. Smol. Paleoecological Environmental Assessment and Research Laboratory (PEARL), Department of Biology, Queen's University, Kingston, ON K7L 3N6, Canada.

[1]Corresponding author (e-mail: ginnb@biology.queensu.ca).

 doi:10.1139/F07-078 © 2007 NRC Canada

Fig. 1. Map of Nova Scotia (Canada) showing locations of study areas. Inset shows location of Nova Scotia within the context of North America.

crease in these lakes compared with the 1950s (Gorham 1957). Continual monitoring of 72 Nova Scotia target lakes has been undertaken by Environment Canada's Long-Range Transport of Airborne Pollutants (LRTAP) program, which was started in the southwestern portion of the province in 1983, with northern areas being added in 1989 (Clair et al. 2002; Pilgrim et al. 2003). The low pH values recorded in these lakes (along with associated low levels of alkalinity) have remained steady over ~20 years of semiannual monitoring (Clair et al. 2001) despite a ~45% reduction in industrial sulphur emissions due to legislation (Jeffries et al. 2000). The annual sulphate deposition is currently 10.5 kg·ha^{-1}·year^{-1} and has been estimated to range from ~8 to 20 kg·ha^{-1}·year^{-1} since 1980 (Shaw 1979; Clair et al. 1989). Most of the acidifying anions deposited in Nova Scotia are carried into the province by prevailing weather systems that pass over the major industrial areas of North America (northeastern USA; Ohio Valley, USA; as well as Ontario and Quebec in Canada), although local industries around Halifax (oil refineries and power generation stations) are also likely a major local source (Kerekes et al. 1986). In addition, Nova Scotia lakes are influenced by their close proximity to the marine environment (no lake is more than 50 km from salt water) and have substantial natural deposition of marine-sourced salts, chloride, and sulphate (10%–16% of the total sulphate deposited) (Clair et al. 2001, 2002; T.A. Clair, Environment Canada, Atlantic Region, P.O. Box 6227, Sackville, NB E4L 1G6, personal communication). Importantly, many acidic lakes in Nova Scotia have a high humic content from dissolved organic carbon (DOC), which contains naturally sourced weak (humic and fulvic) acids (Shaw 1979; Clair et al. 2002). While these weak acids may provide a natural buffer against pH changes, it is not known how these lakes have changed because of acidic deposition, and consequently the realistic mitigation targets for these systems are unknown. Although it is now well known that

© 2007 NRC Canada

Table 1. Current limnological conditions in 51 low-alkalinity study lakes from Nova Scotia, Canada.

Code	Lake name	Latitude (°N)	Longitude (°W)	Max. depth (m)	pH	Colour	DOC ($mg \cdot L^{-1}$)	Gran alkalinity ($mg \cdot L^{-1}$)
Bridgewater								
NS001	Little Wiles	44.40000	64.64500	6.0	5.6	9.5	1.7	−0.69
NS002	Huey	44.39500	64.73700	1.3	6.0	10.5	2.7	0.11
NS003	Annis	44.32500	64.83500	15.7	6.8	4.3	1.4	2.80
NS004	Matthew	44.32500	64.68700	5.2	5.8	67.3	6.7	0.53
NS005	Hirtle	44.47550	64.75000	5.6	6.1	40.8	5.0	0.48
NS006	Rocky	44.72500	64.72500	8.0	6.1	6.8	1.5	−0.02
NS007	Little Tupper	44.42500	64.96500	7.8	6.2	9.8	2.1	0.49
Cape Breton Highlands National Park								
NS008	Mica Hill	46.81750	60.44030	1.0	5.9	81.0	9.0	1.44
NS009	Warren*	46.41000	60.40000	31.0	6.3	49.0	5.3	1.66
NS010	Cradle	46.73200	60.43520	4.2	5.9	19.5	3.8	0.27
NS011	Branch	46.74820	60.45550	6.5	5.0	86.8	9.2	−1.21
NS012	Lake of Islands	46.74820	60.51230	3.1	5.2	72.3	6.4	−0.71
NS013	Dundas #3	46.72100	60.54500	2.1	5.3	97.5	8.7	−0.13
NS014	Dundas #4*	46.71200	60.53520	2.5	5.2	105.5	11.0	−0.10
NS015	White Hill	46.70750	60.58750	2.0	5.1	82.5	6.2	−0.60
NS016	Gull	46.68500	60.54520	2.0	5.3	69.5	5.9	−0.21
NS017	Indian	46.67520	60.56520	3.0	5.7	62.8	6.0	−0.03
NS018	Two Island	46.65740	60.58400	5.5	5.2	72.3	6.4	−0.71
NS019	Glasgow	40.82750	60.48900	4.5	5.2	61.0	5.1	−0.61
NS020	John Dee	46.81590	60.51000	9.4	6.0	78.0	7.4	1.00
NS021	Long	46.81500	60.48900	1.5	6.6	29.0	3.5	3.33
NS022	Round	46.80950	60.50950	2.0	5.4	76.8	6.6	−0.14
NS023	Deer*	46.77520	60.63560	3.0	6.5	48.8	5.9	2.47
Ecum Secum								
NS024	Halfway Brook	44.94750	62.46500	7.6	5.7	41.3	4.7	0.09
Kejimkujik National Park								
NS025	Grafton	44.38500	65.18500	10.0	6.0	44.0	5.3	0.56
NS025	Cobrielle	44.31750	65.23500	6.3	5.4	11.5	2.5	−0.76
NS027	Pebbleloggitch*	44.30250	65.35100	2.5	4.5	94.5	8.8	−1.90
NS028	Peskowesk*	44.32500	65.30000	13.0	4.9	45.5	5.4	−1.14
NS029	Big Dam (West)	44.45870	65.28700	5.6	5.1	95.0	8.3	−0.64
NS030	Big Dam (East)	44.45000	65.26500	4.5	6.1	15.3	4.8	0.42
NS031	Frozen Ocean*	44.45000	65.35200	7.6	4.9	103.5	9.0	−0.78
NS032	Channel	44.43560	65.31200	1.8	4.8	121.5	9.9	−1.14
NS033	Peskawa*	44.32500	65.37500	9.0	4.7	72.5	7.1	−1.32
NS034	Beaverskin	44.31240	65.33510	6.3	5.5	10.8	2.1	−0.57
NS035	Mountain	44.32540	65.26500	14.8	5.3	25.3	3.4	−0.55
NS036	Upper Silver	44.27500	65.25400	5.8	6.1	12.8	2.9	0.43
NS037	Back	44.29250	65.27500	5.8	5.5	26.0	3.6	−0.34
NS038	Loon	44.33500	65.18500	8.2	5.1	77.0	6.9	−0.59
NS039	Kejimkujik	44.37500	65.25000	19.2	4.9	110.3	9.0	−0.98
Yarmouth								
NS040	Trefry	43.83010	66.04750	12.4	6.6	26.0	5.4	2.82
NS041	George	44.00000	66.05000	8.5	5.9	11.3	3.2	−0.01
NS042	Brenton	43.96000	66.07500	3.7	5.1	163.3	13.6	0.07
NS043	Killams	44.00000	66.08000	1.5	6.2	12.0	3.8	0.54
NS044	Allens	43.95000	66.15000	10.0	6.6	11.3	4.4	5.88
NS045	Churchills	43.99050	66.14950	6.0	6.8	12.5	4.3	7.08
NS046	Darlings	43.96000	66.13000	4.1	6.4	28.8	5.9	3.93
NS047	Cedar	44.02510	66.12540	4.2	6.5	20.8	5.4	1.99
NS048	Bird	43.97500	65.95000	5.2	6.7	6.8	2.7	2.86
NS049	Jesse	44.03200	66.01000	5.7	6.3	23.5	5.0	1.25
NS050	Tedford	44.10000	66.02000	4.3	6.4	5.3	3.1	1.22
NS051	Lower Cornings*	44.05000	66.08000	3.8	6.0	27.3	4.9	0.50

Note: An asterisk (*) denotes lake used in core reproducibility study. Water chemistry data courtesy of T.A. Clair (Environment Canada, Atlantic Region, P.O. Box 6227, Sackville, NB E4L 1G6, unpublished data). DOC, dissolved organic carbon; Gran alkalinity, the equivalent sum of titratable bases to the Gran procedure end point. Colour is expressed in relative units.

© 2007 NRC Canada

Table 2. Net changes in diatom-inferred pH among regions of Nova Scotia and three diatom-based inference models.

Core sections (cm)	Model	Kejimkujik National Park	Yarmouth	Bridgewater	Cape Breton Highlands National Park	Ecum Secum
0–15	WA	–0.31	+0.02	–0.06	–0.10	–0.35
	ML	–0.38	–0.09	–0.06	–0.10	–0.26
	MAT	–0.22	–0.05	+0.002	–0.02	–0.41
0–25	WA	–0.41	+0.05	–0.01	+0.04	–0.23
	ML	–0.49	–0.01	–0.10	–0.01	–0.25
	MAT	–0.24	–0.001	+0.09	–0.02	–0.17
15–25	WA	–0.01	–0.02	+0.09	+0.17	+0.12
	ML	–0.08	+0.02	+0.08	+0.13	+0.01
	MAT	–0.03	+0.02	+0.08	–0.01	–0.15

Note: WA, weighted averaging; ML, maximum likelihood; MAT, modern analog technique.

many Nova Scotia lakes are currently acidic (Jeffries et al. 2000), it is not clear if these lakes are naturally acidic or have been acidified because of anthropogenic sulphate deposition.

One highly successful method that has been used to track environmental changes in other regions is paleolimnology, where biological indicators preserved in stratigraphic layers of lake sediments are used as an archive to infer past environmental conditions (e.g., Smol 2002; Moos et al. 2005; Salonen et al. 2006). The Paleoecological Investigation of Recent Lake Acidification and Surface Water Acidification Program studies of the 1980s and early 1990s are prime examples of the utility and rigor of these methods (Battarbee et al. 1990; Cumming et al. 1992). Using these methods, remains of biological organisms (e.g., diatoms (Bacillariophyceae)) from deepwater surface-sediment samples (representing the last 1–2 years of sediment accumulation) are identified and calibrated to current water chemistry data (Birks 1998). If the calibration data set covers a large number of lakes that span an important environmental variable (e.g., pH), the optima and, to a lesser extent, the tolerances of different diatom taxa can be estimated. Based on these estimates, environmental conditions based on the abundance of species within an assemblage (e.g., from pre-industrial samples) can be inferred (Birks et al. 1990; Smol 2002). Using these techniques, past long-term environmental changes that have occurred in a lake can be estimated, including estimated background or pre-impact pH, amount of change, timing, and probable cause. This information can be used to estimate critical loads of sulphate deposition and set realistic mitigation goals.

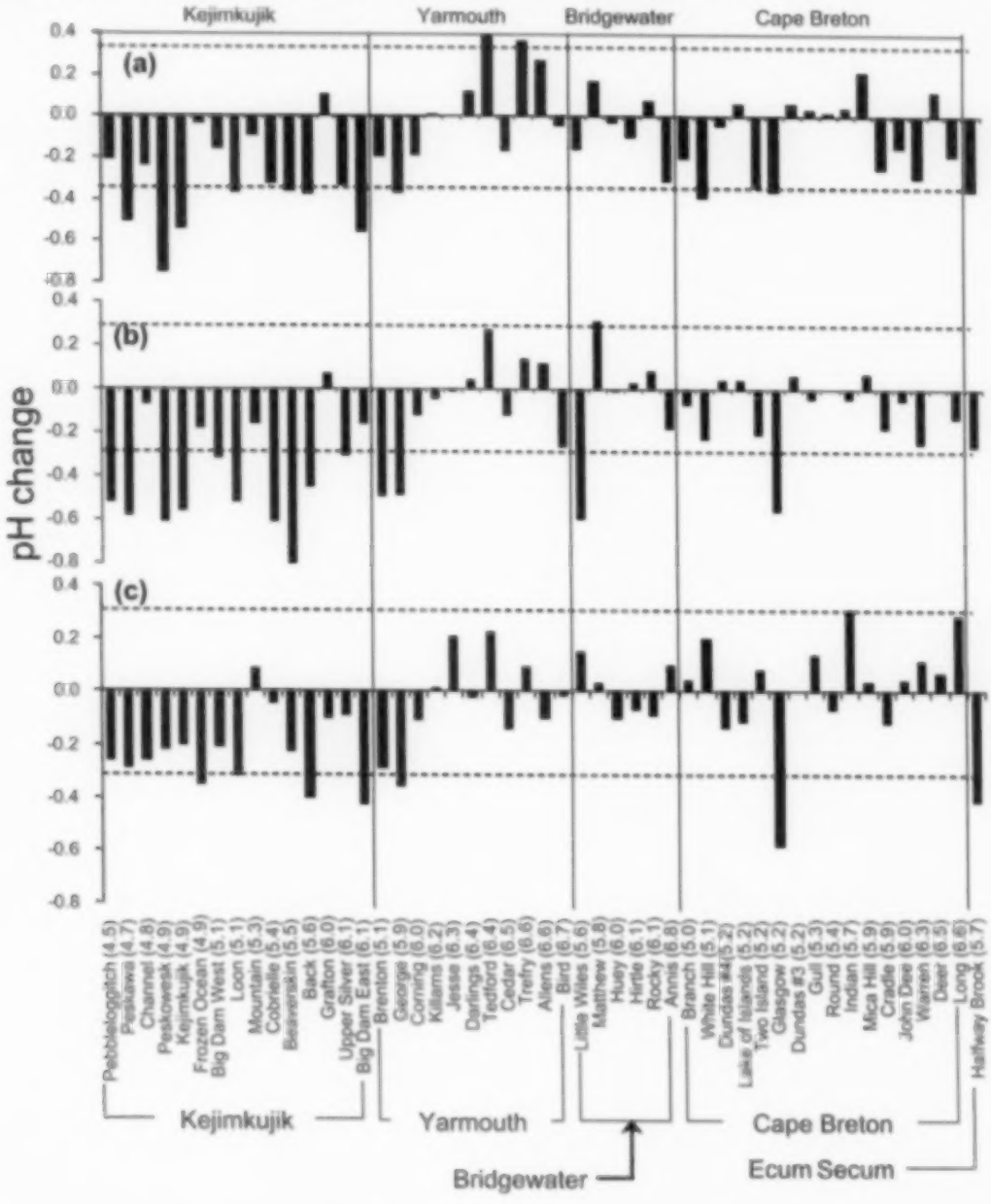
While detailed paleolimnological studies of individual lakes are often required to estimate rates of change, it can be time-consuming to repeat the process for large numbers of lakes, especially when trying to determine how many lakes in a region have been impacted by an environmental stressor. In these cases, a "top–bottom" regional paleolimnological approach (e.g., Cumming et al. 1992) is usually employed to ascertain what lakes have changed and by how much. In this technique, current environmental conditions are inferred based on biological assemblages preserved in the surface sediments (i.e., the top sample). These estimates are compared with the inferred environmental conditions from the biological assemblages present from a layer of the core representing predisturbance conditions (i.e., the bottom sample); in northeastern North America, a depth of 15–20 cm in a sediment core typically corresponds to the 1850s — prior to the onset of acidic deposition (Charles and Smol 1990; Cumming et al. 1994; Smol 2002), which is consistent in Nova Scotia lake sediment cores used in this study that have had chronologies verified using ^{210}Pb geochronology (Ginn et al. 2007*a*).

Few long-term studies have been undertaken to track acidification trends in Nova Scotia. The LRTAP program was started in 1983, well after the onset of acidic deposition. Previous paleolimnological studies (Delorme et al. 1984; Elner and Ray 1987; Duthie 1989) only examined a small number of lakes. Moreover, in these pioneering paleolimnological studies, radiometric dating techniques were not used to estimate sediment chronology, and cores were sectioned and analyzed at relatively coarse intervals (e.g., 1–2 cm intervals representing decadal to multidecadal resolution, in comparison with 0.25 cm resolution, representing subdecadal resolution). More modern methods of statistical inferences of limnological conditions, based on regional calibration data sets and more ecologically sound techniques, have also resulted in more robust pH inference techniques than employed in these earlier studies (e.g., Battarbee et al. 1999; Smol 2002; Ginn et al. 2007*b*).

In this study, diatom-inferred pH changes since pre-industrial times were studied in 51 Nova Scotia lakes that are part of the LRTAP monitoring program (Clair et al. 2001). By using an approach that compared current environmental conditions with pre-industrial environmental conditions (the top–bottom paleolimnological approach), we sought to determine which lakes, in which areas of the province, have been impacted by acidic deposition and how great the changes have been. In addition, this study used a modified top–bottom approach by adding a third sample, if available, to each core, taken from a depth of 25 cm, which corresponds to a time of ~150 years before the onset of acidic deposition (or ~300 years before present). Inclusion of this extra sample enables the comparison of what were the environmental conditions before widespread European settlement (started in 1605 but of small scale until the mid-18th century), thus allowing an assessment of limnological change over a period of minimal or no anthropogenically altered rainfall. The change inferred between these two pre-industrial samples was used as an index of change (or variability) that would be expected in the absence of anthro-

© 2007 NRC Canada

Fig. 2. Change in diatom-inferred pH between surface sediment (current) and 15 cm (~150 years pre-impact) using (*a*) weighted averaging (WA); (*b*) maximum likelihood (ML); and (*c*) modern analog technique (MAT) models. Current measured pH is given in parentheses beside lake name. Churchill and Little Tupper lakes are not shown, as cores were under 15 cm. Broken horizontal lines indicate critical values derived from the natural variation in pH between pre-industrial time periods.



pogenic deposition. Top–bottom paleolimnological studies typically carry a risk that comes with observing only two (or three) samples from a sediment core; there is a possibility that the time of impact by the stressor has been missed completely and that only the pre-impact and recovery periods are being observed (Smol 2002). This study is unusual in that 20 years of water chemistry data are available for the study lakes, which indicate that chemical recovery has not occurred. As such, the lakes are currently as acidic as they were in 1983, near the peak period of acidic deposition (~20 kg·ha^{-1}·year^{-1}) (Clair et al. 2001, 2002).

A second concern with many paleolimnological studies is the ability of one sediment core per lake to accurately represent the environmental changes that occur throughout an

© 2007 NRC Canada

entire water body. While this concern has been addressed in many previous studies (reviewed by Smol 2002), we include a reproducibility component to this study to check the reproducibility of environmental inferences, especially in the larger, more complex lakes.

Materials and methods

Sample collection

Sediment cores were collected between May 2002 and March 2004 from 51 lakes in Nova Scotia that have been monitored as part of the LRTAP program since 1983 (Yarmouth, Bridgewater, and Kejimkujik National Park study areas) or since 1989 (Ecum Secum and Cape Breton Highlands National Park study areas) (Fig. 1, Table 1), using a Glew gravity corer (Glew 1989; Glew et al. 2001) equipped with a 7.6 cm internal diameter, 50 cm long Lexan® tube. Cores were sectioned at 0.25 cm intervals using a Glew extruder (Glew 1988), sections placed in individual Whirl-Pak® bags, and stored at 4 °C. Diatoms were isolated by using a 1:1 molar ratio of HNO_3 and H_2SO_4 as described by Battarbee et al. (2001) and mounted on slides using Naphrax® as a mountant. A minimum of 300 diatom valves were enumerated per sample using a 100× oil-immersion objective (numerical aperture = 1.3) on a Leica DMRB microscope equipped with a Retiga Q-Imaging camera. Diatoms were identified to the lowest taxonomic level possible using Patrick and Reimer (1966, 1975), Krammer and Lange-Bertalot (1991, 1997, 1999, 2000), Round et al. (1990), Camburn and Charles (2000), Fallu et al. (2000), Wehr and Sheath (2002), and other references cited in Ginn et al. (2007*b*).

Inferences of limnological variables

A diatom-based pH inference model, developed from the 494 lake NENA data set from northeastern North America (Pennsylvania to Nova Scotia), with approximately equal distribution across the pH gradient, was used to infer limnological conditions from observed diatom assemblages (Ginn 2006; Ginn et al. 2007*b*). Three models were used to infer pH: weighted averaging with classical deshrinking (WA), maximum likelihood (ML), and modern analog technique (MAT). These models all have very high predictive ability for pH (r^2_{boot} = 0.88–0.89) and are highly significant (Ginn et al. 2007*b*).

To assess the significance of the inferred pH changes, a cut-off value based on the extreme 5% of changes in pH between the two pre-industrial samples was used (i.e., the inferred change between samples located between 25 and 15 cm depth). This resulted in critical cut-off values of ~0.20–0.38 of a pH unit for the WA, ML, and MAT inference models.

Core reproducibility

To assess the reproducibility of coring technique and inference models, eight of our study lakes, selected across the pH and DOC gradients of the Nova Scotia lakes, were replicated by taking three cores from these lakes. All cores were taken from the same basin, although another study (Ginn et al. 2007*c*) further investigated reproducibility in multiple sediment cores from the same and different basins in a large lake (Kejimkujik Lake). Coring methods used were the same as those presented above. Only 21 of 24 replicate cores were of sufficient length to estimate pre-impact inferences; however, all replicate cores were used to compare surface-sediment assemblages.

Table 3. Correlations of inferred pH change among three inference models for three reconstructions (0–15, 0–25, and 15–25 cm).

Interval (cm)	Model	ML	WA	MAT
0–15	ML	1	—	—
	WA	0.79	1	—
	MAT	0.49	0.48	1
0–25	ML	1	—	—
	WA	0.90	1	—
	MAT	0.53	0.42	1
15–25	ML	1	—	—
	WA	0.78	1	—
	MAT	0.24	0.27	1

Note: WA, weighted averaging; ML, maximum likelihood; MAT, modern analog technique. In all cases, $p < 0.01$.

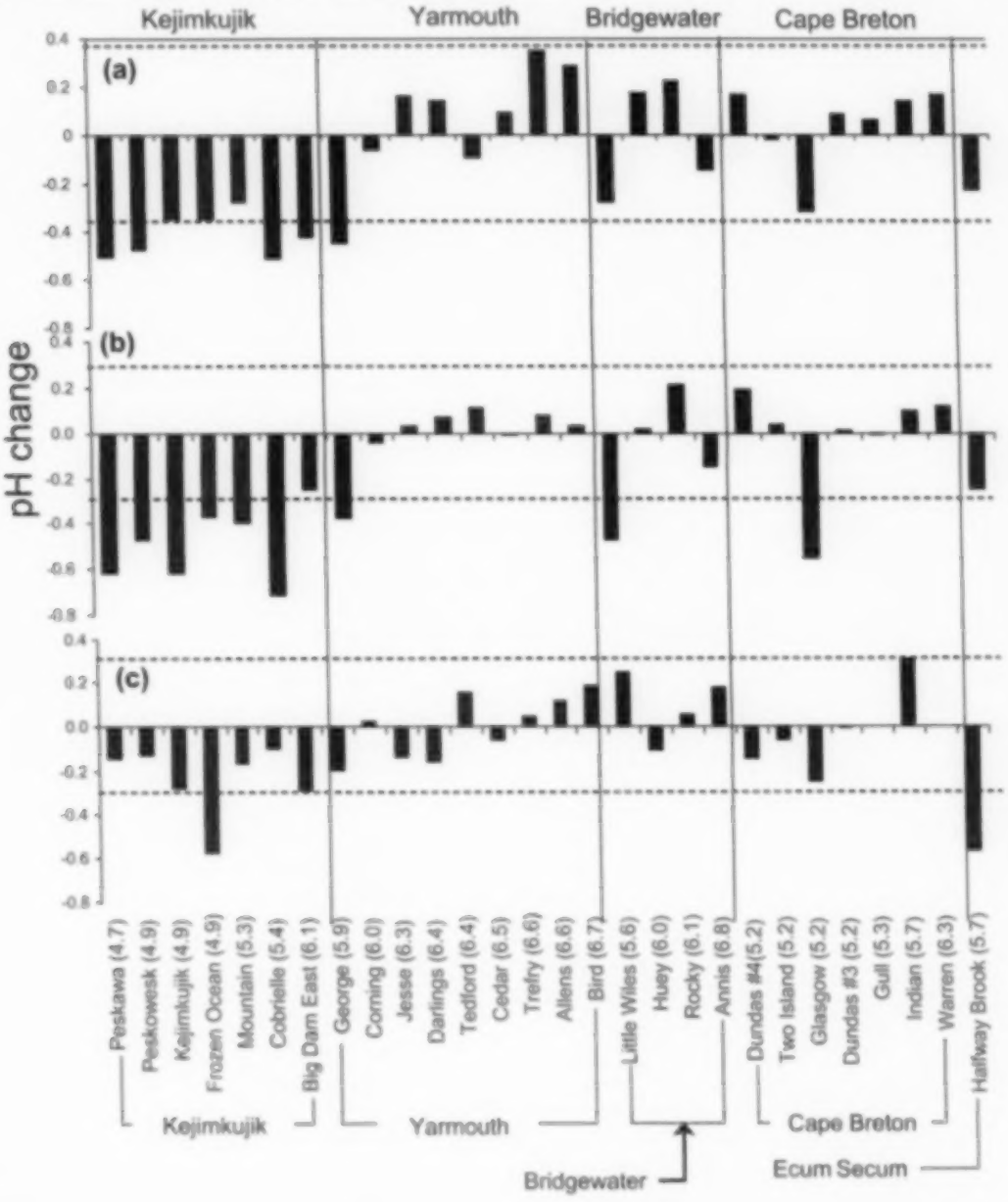
Results

Many of the 51 Nova Scotia LRTAP study lakes record a decrease in diatom-inferred pH between current and pre-impact periods, although the extent of this change varies between regions, with lakes in Kejimkujik National Park showing the greatest pH decrease since the pre-industrial period (Table 2). The study lakes that are currently most acidic have shown the greatest change in pH (ML estimates $r^2 = 0.90$, $p < 0.01$). While similar trends are shown among the three inference models, WA (Fig. 2*a*) and ML (Fig. 2*b*) show the greatest similarity ($r = 0.79$, $p < 0.01$), whereas MAT inferences (Fig. 2*c*) are less similar ($r \sim 0.49$, $p < 0.01$) (Table 3). In addition, MAT, while showing similar overall trends, infers more modest changes in lakewater pH and differs with WA and ML in which individual lakes showed changes. Kejimkujik National Park shows the greatest extent of acidification, with 10 of 15 lakes showing a significant decrease in pH. A significant pH change is shown by the broken line, which is two units of standard deviation away from the average pre-impact pH change for each model (Figs. 3–5). The WA (Fig. 2*a*) and ML (Fig. 2*b*) models show closely related trends with similar mean acidification (WA ~ –0.3 pH unit, ML ~ –0.4 pH unit) (Table 2). MAT (Fig. 2*c*) shows more modest changes (~ –0.2 pH unit) (Table 2). Results in other regions (Yarmouth, Bridgewater, and Cape Breton Highlands National Park) show more variation with less distinct changes in diatom-inferred lakewater pH (Table 2, Fig. 2). While some lakes in Yarmouth and Cape Breton do show significant decreases in lakewater pH, these are not to the same extent and consistency as the lakes in Kejimkujik National Park. The single lake at Ecum Secum, east of Halifax, shows a decrease in diatom-inferred pH using all three inference techniques (Fig. 2*b*).

Similar trends to that observed between the surface and 15 cm (~150 years ago) sections were recorded between the surficial and 25 cm (~300 years ago) core sections (Fig. 3). Again, WA (Fig. 3*a*) and ML (Fig. 3*b*) showed very similar trends ($r = 0.90$, $p < 0.01$). Lakes in Kejimkujik National

© 2007 NRC Canada

Fig. 3. Change in diatom-inferred pH between surface sediment (current) and 25 cm (~300 years pre-impact) using (*a*) weighted averaging (WA); (*b*) maximum likelihood (ML); and (*c*) modern analog technique (MAT) models. Current measured pH is given in parentheses beside lake name. Sample size is smaller because cores were under 25 cm. Broken horizontal lines indicate critical values derived from the natural variation in pH between pre-industrial time periods.

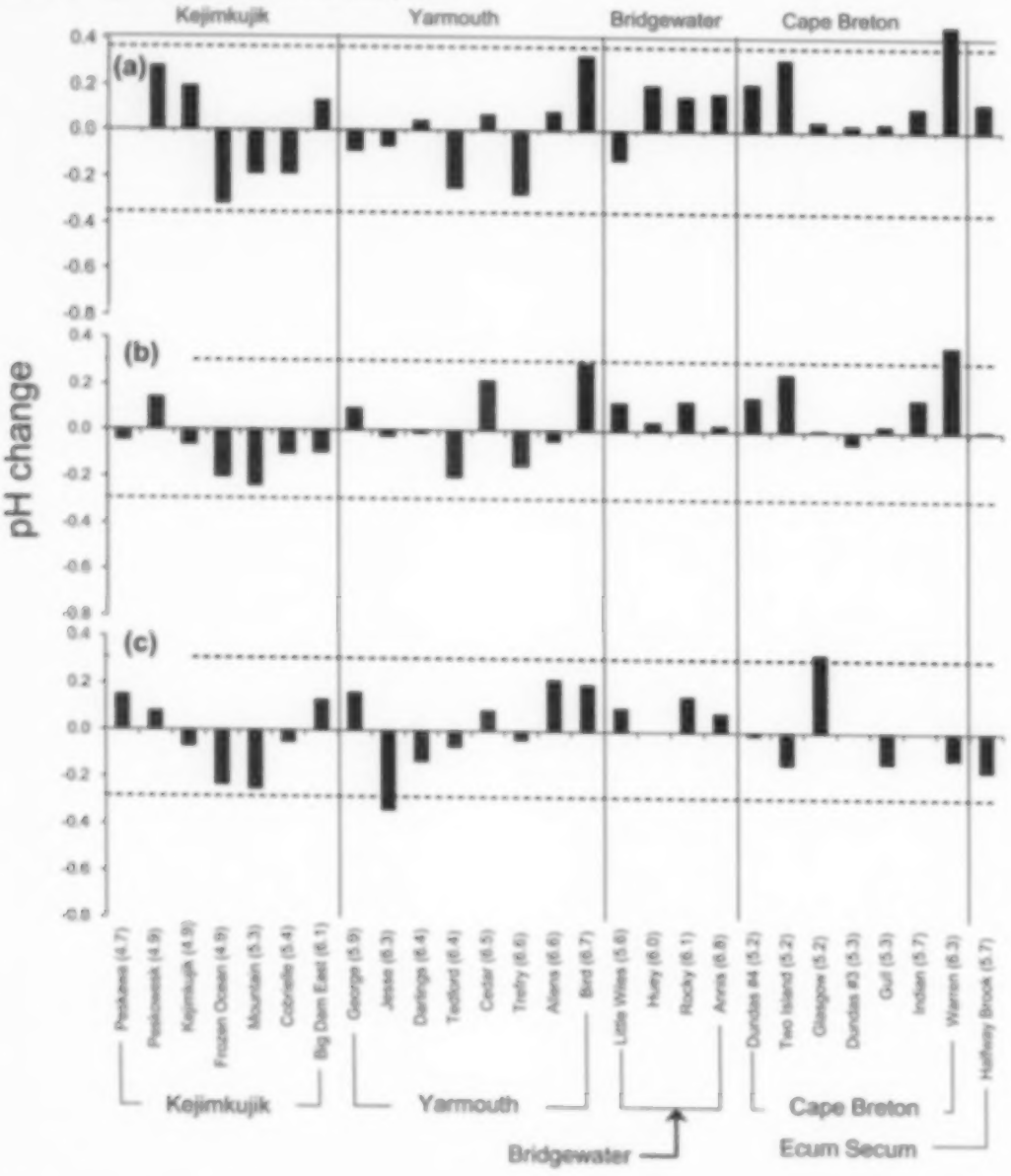


Park showed the greatest decreases in diatom-inferred pH (WA = –0.41 pH unit, ML = –0.49 pH unit (Table 2)), virtually identical to patterns indicated between ~150 years ago pre-industrial and present-day samples (Fig. 2). MAT (Fig. 3*c*) indicated more modest acidification trends (MAT = –0.24 pH unit) (Table 2) in Kejimkujik National Park, although the trends inferred by this model are comparable to the ~1850 to present-day inferred changes (Fig. 3*c*).

Changes between the two pre-industrial pH values (~150 and ~300 years before present) between regions are small and allowed derivation of critical values encompassing 95% of the observations (Table 2, Fig. 4), which was used to set a critical value (two times the standard deviation or 0.2 pH unit) for substantial anthropogenic-caused lakewater pH change. Changes in diatom-inferred pH are negligible in all three models, although WA (Fig. 4*a*) and ML (Fig. 4*b*) are

© 2007 NRC Canada

Fig. 4. Change in diatom-inferred pH between 15 and 25 cm (both ~150 and ~300 years pre-impact) using (*a*) weighted averaging (WA); (*b*) maximum likelihood (ML); and (*c*) modern analog technique (MAT) models. Current measured pH is given in parentheses beside lake name. Sample size is smaller because cores were under 25 cm. Broken horizontal lines indicate critical values derived from the natural variation in pH between pre-industrial time periods.



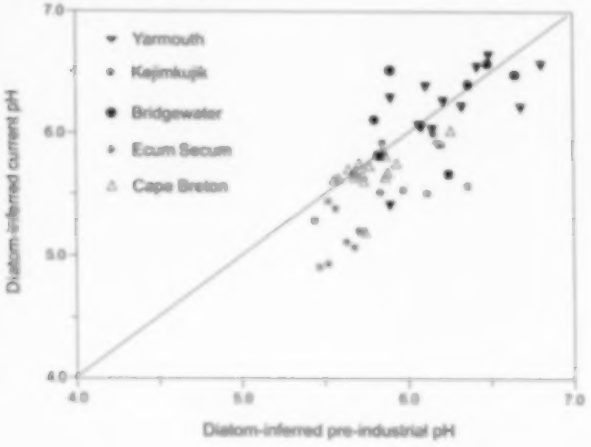
most similar. As noted before, MAT (Fig. 4c) shows different changes from the other models. ML and WA have highly correlated results ($r = 0.78$, $p < 0.01$), whereas MAT is less strongly correlated with these other models ($r \sim 0.24$) (Table 3). Additionally, lakes in Kejimkujik National Park tend to have lower pre-industrial pH values (diatom-inferred pH range 5.4–6.2) in comparison with lakes in Yarmouth and Bridgewater (Fig. 5).

Although the range in pre-industrial pH values between Kejimkujik National Park and Cape Breton overlap, four of the Kejimkujik area lakes have the lowest pre-industrial inferred pH values (Fig. 5). Lakes from Kejimkujik National Park showed a larger decline in pH in comparison with the minimal changes noted in the more northern Cape Breton lakes (Table 2, Fig. 5).

With respect to core reproducibility, the tops of each core from eight lakes were highly similar as were the bottom (15 cm) samples (Fig. 6). The pH changes shown in each sample are highly reproducible and mirror the acidification trends in detailed analysis of sediment cores from

© 2007 NRC Canada

Fig. 5. Summary of acidification in 51 Nova Scotia lakes: lakes with the lowest pre-impact pH in a high sulphate deposition area (Kejimkujik National Park) acidified most, whereas other areas remain relatively unchanged.



Kejimkujik National Park and Cape Breton Highlands National Park (Ginn et al. 2007*a*). Inferences using WA and MAT showed identical trends with respect to reproducibility (Ginn 2006).

Discussion

Many lakes in Nova Scotia have been impacted by acidic deposition between pre-industrial and present-day time periods. While the extent of this acidification trend varies across the province, lakes in the southwest have been much more severely impacted than lakes in northern areas. This decreasing acidification impact could be due to a decrease in sulphate deposition (highest in the southwest, lowest in the north) and a similar trend in the pH of precipitation, which is lowest in southwest Nova Scotia (pH = 4.5–4.7) and highest in Cape Breton (pH = 5.0) (Kerekes et al. 1986; Underwood et al. 1987). Early records of precipitation pH in Nova Scotia date from the 1950s (pH = ~5.0), and while they date before the suspected peak sulphate deposition (early 1970s), the recorded pH is still likely lower than pre-industrial values. In addition, lakes in Kejimkujik National Park have lower inferred pre-industrial pH values (mean pH: 5.8, range: 5.4–6.2) in comparison with lakes in the Yarmouth (mean: 6.3, range: 5.9–6.8) or Bridgewater (mean pH: 6.2, range: 5.8–6.7) study areas. While all southwestern study areas have similar acid-sensitive geology, thin soils, and sulphate loading, the lower pre-industrial inferred pH for lakes in Kejimkujik National Park might account for the greater impact from acidic deposition. This finding is similar to that by Smol et al. (1998), which showed that Adirondack lakes (pre-industrial pH range: 5.0–7.0) acidified under a lower deposition (in comparison with lakes in Sudbury with pre-industrial pH range: 6.0–7.8). In addition, lakes in the Adirondack region have not responded to decreases in sulphate loading. The lakes of Kejimkujik National Park would be in the same category as the most acid-sensitive Adirondack lakes (Cumming et al. 1994; Smol et al. 1998). In addition, watershed development in Yarmouth and Bridgewater (i.e., logging, farming, or residential), compared with Kejimkujik National Park (protected since 1967), may have added buffering capacity by release of base cations or a masking of the acidification signal by changes in nutrient status or some other environmental stressor (Davis et al. 1994). Grafton Lake in Kejimkujik National Park has had disturbances in its watershed (dam construction and removal, logging in watershed outside park boundaries) and is one lake that two inference models showed had not acidified; in fact, there has been a slight alkalinization trend. The inferred pre-industrial lakewater pH of Grafton Lake was 5.9, relatively high in comparison with other Kejimkujik area lakes, and suggests the lake would be better buffered against acidic deposition in comparison with lakes with a low background pH value.

Lakes in Cape Breton Highlands National Park have a diatom-inferred pre-industrial pH (mean pH: 5.8, range: 5.6–6.3) similar to that of Kejimkujik National Park, yet there is little inferred acidification in Cape Breton Highlands. A core analysis with decadal scale resolution of six Cape Breton lakes (White Hill, Dundas #4, Lake of Islands, Cradle, Warren, and Deer) did not show changes in diatom assemblages or inferred pH that are consistent with acidification (Ginn et al. 2007*a*). As sulphate deposition is less (and precipitation pH is higher) in Cape Breton (Clair et al. 2002) compared with southwestern Nova Scotia, these lakes probably did not receive enough sulphate to result in a substantial pH decline in these lakes.

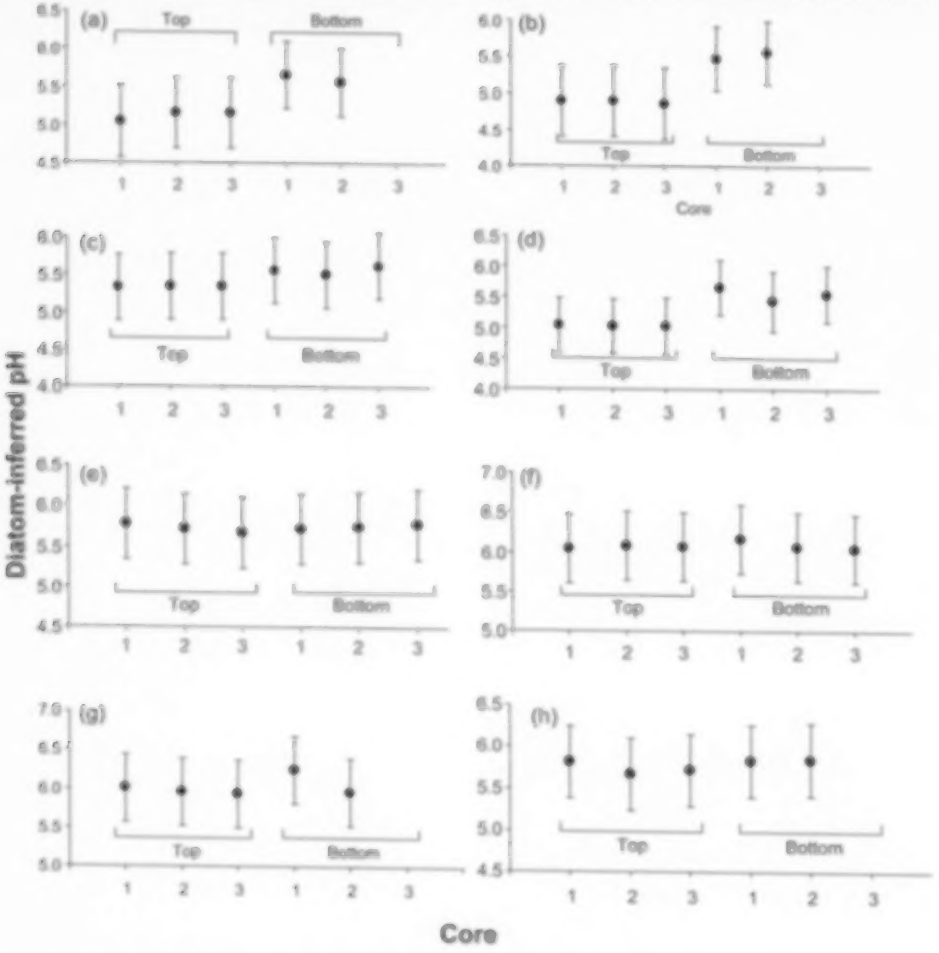
The decreasing pattern of sulphate deposition across Nova Scotia is due to a gradient of sulphate deposition, with ~90% of sulphate arriving through long-range transport on air currents that pass over the major industrial centres of North America (i.e., Ohio Valley and northeastern USA, Ontario and Quebec, Canada) (Shaw 1979; Thomas 1983; Clair et al. 1989). An exception to this trend is the lake in the Ecum Secum study area, which is downwind of sulphur-emitting industries in the city of Halifax (power generation plants, oil refineries). While one lake is not sufficient to draw important conclusions, the lakewater pH decrease we recorded may be expected, as this area receives lower pH precipitation than locations an equal distance upwind of Halifax (Underwood et al. 1987).

Inferences of pH change between surface sediments and the 25 cm core intervals show trends similar to the 0–15 cm analysis discussed above. While the sample size is smaller (not all lake cores were over 25 cm in length), the trends observed show pH decreases in all Kejimkujik National Park lakes; less obvious trends in Yarmouth, Bridgewater, and Cape Breton; and a pH decrease in the single Ecum Secum site east of Halifax.

The inclusion of a second pre-industrial sample (~300 years before present) into this paleolimnological analysis enables the study of pH change over a similar time scale (~150 years) within the pre-industrial period. Using the samples at 25 cm, the natural variation in pH over ~150 years was estimated and used to set a critical pH threshold for significant change. Pre-industrial trends in these lakes show a more ambiguous pattern with much smaller pH changes in comparison with the post-industrial period. The critical thresholds used were set using a conservative estimate of twice the standard deviation of pH change in each model. These values (WA: 0.38 pH unit; ML: 0.20 pH unit; MAT:

© 2007 NRC Canada

Fig. 6. Reproducibility of maximum likelihood (ML) pH inferences, with vertical standard error bars among three cores (1–3) from eight lakes (current measured pH in parentheses): (*a*) Pebbleloggitch Lake (4.5); (*b*) Peskawa Lake (4.7); (*c*) Frozen Ocean Lake (4.9); (*d*) Peskowesk Lake (4.9); (*e*) Dundas #4 Lake (5.2); (*f*) Lower Cornings Lake (6.0); (*g*) Warren Lake (6.3); (*h*) Deer Lake (6.5).



0.31 pH unit) are similar but slightly smaller than the root mean squared error of prediction (RMSEP) of the inference models (WA: 0.43; ML: 0.45; MAT: 0.46 (Ginn et al. 2007*b*)), which have been used to set critical limits in other studies (e.g., Cumming et al. 1992, 1994).

Of the three inference models used, we believe that the ML and WA approaches provide the best estimate of the actual pH changes that are taking place in the systems. In a decadal-scale detailed analysis of Nova Scotia sediment cores, ML correlated most closely ($r = 0.9$) to the main direction of variation determined using PCA axis 1 scores (Ginn et al. 2007*b*). In addition, WA ($r = 0.8$, $p < 0.01$) was concluded to provide good estimates of ML results, as was shown by the close correlation of WA and ML inferences ($r = 0.9$) in this study. MAT, although similar to WA and ML ($r = 0.5$), was found to be less reflective of biological changes in detailed analyses (i.e., the correlation of MAT inferences with PCA axis 1 scores: $r = 0.6$, $p < 0.01$) (Ginn et al. 2007*a*). Additionally, the many subjective decisions necessary when using MAT approaches make this technique less desirable (Ginn et al. 2007*b*), as does its sensitivity to the distribution of lakes in a data set (Ginn et al. 2007*b*). While MAT does show acidification-related trends, the inferred pH changes are more modest than the other models. This may be due to a lack of analog assemblages at more acidic pH values, which reflect the current environmental conditions of Nova Scotia lakes. Consequently, as additional analogs are chosen, the similarities between the samples decrease and pull the inferences away from the lower pH values (Overpeck et al. 1985; Pflaumann et al. 1990; Ginn et al. 2007*b*). The advantage of ML and WA is that they use all of the environmental and species data to generate an inference model, whereas MAT uses closest matches from a similarity index and consequently results in "jumpier" and more noisy inferences (Ginn et al. 2007*a*, 2007*b*).

Inferences from both the top and bottom sediment samples of the cores from eight reproducibility lakes show good agreement with inferences within tops (0–0.5 cm interval) and bottoms (15.0–15.25 cm interval) being less variable in comparison with the variability between top and bottom samples

© 2007 NRC Canada

in the lakes that were inferred to acidify. Pebbleloggitch, Peskawa, and Peskowesk lakes were inferred to have acidified, whereas the other lakes underwent lakewater pH changes less than the critical value. This same conclusion emerges when examining the results from multiple cores and are similar to reproducibility data from other paleolimnological studies (reviewed by Smol 2002), which found that multiple cores taken from the deep basins of a lake often show highly correlated trends in both percentage assemblage changes and in related inferences, especially when cores are collected from a lake's deep basins where variability in diatom assemblages is low (Hall and Smol 1996). In addition, a study on the multiple basins from a single, large lake (Kejimkujik Lake) by Ginn et al. (2007*c*) showed strong and highly correlated trends in respect to both the amplitude and timing of biotic and diatom-inferred pH changes.

In conclusion, acidification has affected Nova Scotia lakes situated on nonbuffering bedrock; however, the degree of diatom-inferred pH change has been significantly different for various regions of the province. Of the three study areas in southwestern Nova Scotia that have relatively high sulphate deposition, Kejimkujik National Park had the largest number of lakes with pH decreases (10 of 15 lakes are significantly more acidic now compared with ~150 years ago), while only three lakes from the Yarmouth and Bridgewater regions significantly decreased in lakewater pH. Despite the sensitivity to acidification in these three areas, lakes in Kejimkujik National Park had a lower pre-industrial pH, and as such probably acidified earlier and under less sulphate loading than Yarmouth and Bridgewater lakes, similar to that observed in the Adirondack region lakes (Smol et al. 1998). In addition, development in the watersheds of Yarmouth and Bridgewater lakes and higher pre-industrial pH values likely resulted in greater buffering against pH changes (Davis et al. 1994). While only 1 of 17 lakes in Cape Breton significantly acidified, background pH values were primarily between 5.5 and 6.0, just slightly higher than lakes in Kejimkujik National Park. Cape Breton lakes, despite their potential sensitivity, are located in a lower sulphate deposition region and, based on our data, did not reach the critical loading of sulphate deposition resulting in acidification.

Acknowledgements

This study would not have been possible without the assistance of Tom Clair, Environment Canada – Atlantic Region, for providing water chemistry data for Nova Scotia; Parks Canada staff at Kejimkujik National Park and Cape Breton Highlands National Park; and M. Rate, C. Chan, B. Keddy, and B. Keatley for field assistance. This study was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic Grant to JPS, BFC, and Peter Dillon and also by an Ontario Graduate Scholarship and R.S. McLaughlin Fellowship to BKG.

References

Battarbee, R.W., Mason, J., Renberg, I., and Talling, J.F. 1990. Palaeolimnology and lake acidification. The Royal Society, London.

Battarbee, R.W., Charles, D.F., Dixit, S.S., and Renberg, I. 1999. Diatoms as inidicators of surface water acidity. *In* The diatoms: applications for the environmental and earth sciences. *Edited by* E.F. Stoermer and J.P. Smol. Cambridge University Press, Cambridge, UK. pp. 85–127.

Battarbee, R.W., Jones, V.J., Flower, R.J., Cameron, N.G., Bennion, H., Carvalho, L., and Juggins, S. 2001. Diatoms. *In* Tracking environmental change using lake sediments. Vol. 3. Terrestrial, algal, and siliceous indicators. *Edited by* J.P. Smol, H.J.B. Birks, and W.M. Last. Kluwer Academic Press, Dordrecht, The Netherlands. pp. 155–202.

Birks, H.J.B. 1998. D.G. Frey & E.S. Deevey Review #1: Numerical tools in palaeolimnology — progress, potentialities, and problems. J. Paleolimnol. **20**: 307–332.

Camburn, K.E., and Charles, D.F. 2000. Diatoms of low-alkalinity lakes in the northeastern United States. Acad. Nat. Sci. Philadelphia Spec. Publ. 18.

Charles, D.F., and Smol, J.P. 1990. The PIRLA II project: regional assessment of lake acidification trends. Verh. Int. Ver. Limnol. **24**: 474–480.

Clair, T.A., Schwarcz, H.P., and Kramer, J.R. 1989. The origins of sulfur in waters from four Nova Scotia basins, Canada. Appl. Geochem. **4**: 93–98.

Clair, T.A., Pollock, T., Brun, G., Ouellet, A., and Lockerbie, D. 2001. Environment Canada's acid precipitation monitoring networks in Atlantic Canada. Environment Canada, Ottawa. Occas. Rep. No. 16.

Clair, T.A., Ehrman, J.M., Ouellet, A.J., Brun, G., Lockerbie, D., and Ro, C.-U. 2002. Changes in freshwater acidification trends in Canada's Atlantic provinces: 1983–1997. Water Air Soil Pollut. **138**: 335–354.

Cumming, B.F., Smol, J.P., Kingston, J.C., Charles, D.F., Birks, H.J.B., Camburn, K.E., Dixit, S.S., Uutala, A.J., and Selle, A.R. 1992. How much acidification has occurred in Adirondack region lakes (New York, USA) since pre-industrial times? Can. J. Fish. Aquat. Sci. **49**: 128–141.

Cumming, B.F., Davey, K.A., Smol, J.P., and Birks, H.J.B. 1994. When did acid-sensitive Adirondack lakes (New York, USA) begin to acidify and are they still acidifying? Can. J. Fish. Aquat. Sci. **51**: 1550–1568.

Davis, R.B., Anderson, D.S., Norton, S.A., Ford, J., Sweets, P.R., and Kahl, J.S. 1994. Sedimented diatoms in northern New England lakes and their use as pH and alkalinity indicators. Can. J. Fish. Aquat. Sci. **51**: 1855–1876

Delorme, L.D., Esterby, S.R., and Duthie, H. 1984. Prehistoric pH trends in Kejimkujik Lake, Nova Scotia. Int. Rev. Hydrobiol. **69**: 41–55.

Duthie, H. 1989. Diatom-inferred pH history of Kejimkujik Lake, Nova Scotia: a reinterpretation. Water Air Soil Pollut. **46**: 317–322.

Elner, J.K., and Ray, S. 1987. pH profiles from diatom stratigraphies in sediment cores of selected lakes of New Brunswick and Nova Scotia, Canada. Water Air Soil Pollut. **32**: 17–29.

Fallu, M.-A., Allaire, N., and Pienitz, R. 2000. Freshwater diatoms from northern Québec and Labrador (Canada). Bibl. Diatomol. **45**: 1–200.

Ginn, B.K. 2006. Assessment of surface-water acidification using diatoms as paleoecological indicators in low alkalinity lakes in Nova Scotia (Canada) with a focus on lakes in Kejimkujik and Cape Breton Highlands National Parks. Ph.D. dissertation, Department of Biology, Queen's University, Kingston, Ont.

Ginn, B.K., Cumming, B.F., and Smol, J.P. 2007*a*. Long-term acidification trends in high- and low-sulphate deposition regions in Nova Scotia, Canada. Hydrobiologia, **586**: 261–275.

© 2007 NRC Canada

Ginn, B.K., Cumming, B.F., and Smol, J.P. 2007*b*. Diatom-based environmental inferences and model comparisons from 494 northeastern North American lakes. J. Phycol. **43**. In press. doi:10.1111/j.1529-8817.2007.00363.x.

Ginn, B.K., Stewart, L.J., Cumming, B.F., and Smol, J.P. 2007*c*. Surface-water acidification and reproducibility of sediment cores from Kejimkujik Lake, Nova Scotia, Canada. Water Air Soil Pollut. **183**: 15–24.

Glew, J. 1988. A portable extruding device for close interval sectioning of unconsolidated core samples. J. Paleolimnol. **1**: 235–239.

Glew, J. 1989. A new trigger mechanism for sediment samplers. J. Paleolimnol. **2**: 241–243.

Glew, J.R.G., Smol, J.P., and Last, W.M. 2001. Sediment core collection and extrusion. *In* Tracking environmental change using lake sediments. Vol. 1. Basin analysis, coring, and chronological techniques. *Edited by* W.M. Last and J.P. Smol. Kluwer Academic Press, Dordrecht, The Netherlands. pp. 73–105.

Gorham, E. 1957. The chemical composition of lakewaters in Halifax County, Nova Scotia. Limnol. Oceanogr. **2**: 12–21.

Hall, R.I., and Smol, J.P. 1996. Paleolimnological assessment of long-term water-quality changes in south-central Ontario lakes affected by cottage development and acidification. Can. J. Fish. Aquat. Sci. **53**: 1–17.

Jeffries, D.S., Lam, D.C.L., Wong, I., and Moran, M.D. 2000. Assessment of changes in lake pH in southeastern Canada arising from present levels and expected reductions in acidic deposition. Can. J. Fish. Aquat. Sci. **57**(Suppl. 2): 40–49.

Kerekes, J., Howell, G., Beauchamp, S., and Pollock, T. 1982. Characterization of three lake basins sensitive to acid precipitation in central Nova Scotia (June 1979 to May 1980). Int. Rev. Hydrobiol. **67**: 679–694.

Kerekes, J., Beauchamp, S., Tordon, R., Tremblay, C., and Pollock, T. 1986. Organic versus anthropogenic acidity in tributaries of the Kejimkujik watersheds in western Nova Scotia. Water Air Soil Pollut. **31**: 165–173.

Krammer, K., and Lange-Bertalot, H. 1991. Bacillariophyceae 4. Teil: Achnanthaceae, Kritische Ergänzungen *zu* Navicula (Lineolatae) und Gomphonema Gesamtliteraturverzeichnis Teil 1–4. *In* Süßwasserflora von Mitteleuropa 2/4. *Edited by* H. Ettel, J. Gerloff, H. Heynig, and D. Mollenhauer. Spektrum Akademischer Verlag, Berlin.

Krammer, K., and Lange-Bertalot, H. 1997. Bacillariophyceae 2. Teil: Bacillariaceae, Epithemiaceae, Surirellaceae. *In* Süßwasserflora von Mitteleuropa 2/2. *Edited by* H. Ettel, J. Gerloff, H. Heynig, and D. Mollenhauer. Spektrum Akademischer Verlag, Berlin.

Krammer, K., and Lange-Bertalot, H. 1999. Bacillariophyceae 1. Teil: Naviculaceae. *In* Süßwasserflora von Mitteleuropa 2/1. *Edited by* H. Ettel, J. Gerloff, H. Heynig, and D. Mollenhauer. Spektrum Akademischer Verlag, Berlin.

Krammer, K., and Lange-Bertalot, H. 2000. Bacillariophyceae 3. Teil: Centrales, Fragilariaceae, Eunotiaceae. *In* Süßwasserflora von Mitteleuropa 2/3. *Edited by* H. Ettel, J. Gerloff, H. Heynig, and D. Mollenhauer. Spektrum Akademischer Verlag, Berlin.

Lusis, M.A., and Philips, M.L. 1990. The 1990 Canadian long-range transport of air pollutants and acid deposition assessment report. Part 3. Atmospheric sciences. Federal/Provincial Research and Monitoring Coordinating Committee, Ottawa.

Moos, M.T., Laird, K.R., and Cumming, B.F. 2005. Diatom assemblages and water depth in Lake 239 (Experimental Lakes Area, Ontario): implications for paleoclimatic studies. J. Paleolimnol. **34**: 217–227.

Morrison, H.A. 2004. Canadian acid deposition science assessment. Environment Canada, Ottawa.

Overpeck, J.T., Webb, T., and Prentice, I.C. 1985. Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. Quat. Res. **23**: 87–108.

Patrick, R., and Reimer, C.W. 1966. The diatoms of the United States, exclusive of Alaska and Hawaii. Vol. 1. Fragilariaceae, Eunotiaceae, Achnanthaceae, Naviculaceae. Academy of Natural Sciences of Philadelphia Monograph No. 13.

Patrick, R., and Reimer, C.W. 1975. The diatoms of the United States, exclusive of Alaska and Hawaii. Vol. 2. Part 1. Entomoneidaceae, Cymbellaceae, Gomphonemaceae, Epithemiaceae. Academy of Natural Sciences of Philadelphia Monograph No. 13.

Pflaumann, U., Duprat, J., Pujol, C., and Labeyrie, L.D. 1990. SIMMAX: a modern analog technique to deduce Atlantic sea surface temperatures from planktonic foraminifera in deep-sea sediments. Paleoceanography, **11**: 15–35.

Pilgrim, W., Clair, T.A., Choute, J., and Hughes, R. 2003. Changes in acid precipitation-related water chemistry of lakes from southwestern New Brunswick, Canada, 1986–2001. Environ. Monit. Assess. **88**: 39–52.

Round, F.E., Crawford, R.M., and Mann, D.G. 1990. The diatoms: biology and morphology of the genera. Cambridge University Press, Cambridge.

Salonen, V., Tuovinen, N., and Valpola, S. 2006. History of mine drainage impact on Lake Orijäryi algal communities, SW Finalnd. J. Paleolimnol. **35**: 289–303.

Shaw, R.W. 1979. Acid precipitation in Atlantic Canada. Environ. Sci. Technol. **13**: 407–411.

Smol, J.P. 2002. Pollution of lakes and rivers: a paleoenvironmental perspective. Arnold Publishers, London.

Smol, J.P., Cumming, B.F., Dixit, A.S., and Dixit, S.S. 1998. Tracking recovery patterns in acidified lakes: a paleolimnological perspective. Restor. Ecol. **6**: 318–326.

Thomas, M.L.H. (*Editor*). 1983. Meteorology. *In* Marine and coastal systems of the Quoddy region. Can. Spec. Publ. Fish. Aquat. Sci. 64. pp. 5–8.

Underwood, J.K., Vaughn, H.H., Ogden, J.G., and Mann, C.G. 1982. Acidification of Nova Scotia lakes. II. Ionic balances in dilute waters. Nova Scotia Department of the Environment, Halifax, N.S.

Underwood, J.K., Ogden, J.G., Kerekes, J.J., and Vaughn, H.H. 1987. Acidification of Nova Scotia lakes. III. Atmospheric deposition of SO_2 and NO_2 and effects on urban and rural lakes. Water Air Soil Pollut. **32**: 77–88.

Watt, W.D. 1987. A summary of the impact of acid rain on Atlantic salmon (*Salmo salar*) in Canada. Water Air Soil Pollut. **35**: 27–35.

Watt, W.D., Scott, D., and Ray, S. 1979. Acidification and other chemical changes in Halifax County lakes after 21 years. Limnol. Oceanogr. **24**: 1154–1161.

Watt, W.D., Scott, C.D., Zamora, P.J., and White, W.J. 2000. Acid toxicity levels in Nova Scotian rivers have not declined in synchrony with the decline in sulfate levels. Water Air Soil Pollut. **118**: 203–229.

Wehr, J.D., and Sheath, R.G. 2002. Freshwater algae of North America. Academic Press, New York.

Wiltshire, J.F., and Machell, J.R. 1981. A study of acidification in sixteen lakes in mainland Nova Scotia and southern New Brunswick. Environmental Protection Service, Halifax, N.S.

© 2007 NRC Canada

State–space model for light-based tracking of marine animals

Anders Nielsen and John R. Sibert

Abstract: A coherent model is presented to estimate the most probable track of geographic positions directly from a series of light measurements. The model estimates two geographic positions per day, without reducing the daily light data to two threshold crossing times, its covariance structure is designed to handle high correlations due to for instance local weather conditions, and it can estimate the yearly pattern in latitudinal precision by propagating the data uncertainties through the geolocation process. The model is applied to one mooring study, one GPS drifter buoy study, and numerous simulated cases. The simulations are performed with realistic assumptions about the relationship between solar altitude and light and with realistic uncertainty parameters (all taken from real data). The simulations showed that all model parameters were identifiable, and that all tracks could be reconstructed within 1° or 2° latitude and 0.5° or 1° longitude. The mooring and drifter buoy data showed that the tracks could be reliably estimated, even in cases where the other methods had completely failed.

Résumé : Nous présentons un modèle cohérent pour estimer directement le tracé le plus probable des positions géographiques à partir d'une séries de mesures de lumière. Le modèle estime deux positions géographiques par jour, sans réduire les données de lumière journalières à deux périodes de traversée du seuil ; il possède une structure de covariance conçue pour tenir compte des fortes corrélations dues, par exemple, aux conditions climatiques locales ; il peut aussi estimer le patron annuel de précision des latitudes en faisant passer les incertitudes des données dans le processus de géopositionnement. Nous utilisons le modèle sur une étude à bouée fixe, une étude à bouée dérivante munie d'un GPS et plusieurs cas de simulation. Les simulations comportent des présuppositions réalistes concernant la relation entre l'altitude du soleil et la lumière, ainsi que des paramètres d'incertitude réalistes (tous tirés de données réelles). Les simulations montrent que tous les paramètres du modèle sont identifiables et que tous les tracés peuvent être reconstitués avec une précision de 1° ou 2° latitude et de 0,5° ou 1° longitude. Les données obtenues des bouées fixes et dérivantes montrent qu'il est possible d'estimer les tracés de façon fiable, même dans les cas où les autres méthodes ont été complètement incapables de le faire.

[Traduit par la Rédaction]

Introduction

Archival tags, also referred to as data storage tags, are electronic devices designed to be attached externally or surgically implanted into a live animal to record measurements of the animal's environment. For terrestrial animals and for marine animals that stay at the surface for long periods, satellite-based methods can be used to give a very accurate track. Marine species live in salt water where radio waves cannot penetrate, hence these direct position measurements are impossible. For these species, measurements of light, temperature, and pressure are typically stored in the tags, in the hope that reliable geographic positions (geolocations) can be estimated from these data downloaded from recovered archival tags.

Fisheries research organizations, universities, and governments spend large amounts of funds on development, procurement, and deployment of archival tags. This is not because the technology is fully developed, but simply because there really is no alternative. Directly observing these animals in situ is only possible for very short periods. Archival tags can stay on for years and potentially provide valuable information about after-release survival, spawning areas, habitat preferences, and migration corridors (e.g. Arnold and Dewar 2001; Block 2005).

Algorithms for estimating geolocation based on light and pressure have at least a 20-year history (Smith and Goodman 1986; Musyl et al. 2001; Ekstrom 2004). The algorithms are best described as threshold algorithms. From the depth-corrected time series of light measurements, they determine the time when a certain light threshold is crossed. This threshold is assumed to correspond to a certain solar altitude (i.e. at two times a day the solar altitude is (assumed) known). Solar altitude is measured in degrees above the horizon (negative below). The longitude can be computed from local noon (the midpoint between these two times) and latitude from the time elapsed between these two times (day length). The last part of the algorithms, computing position from two known times and angles, is unproblematic, based in celestial mechanics, and well described in textbooks on astronomical algorithms (Meeus 1998).

The problematic part is associating a threshold with a certain solar altitude, as the time when the daily light measurements crosses a certain threshold is influenced by multiple factors (e.g. local weather conditions). Cloud cover will advance the

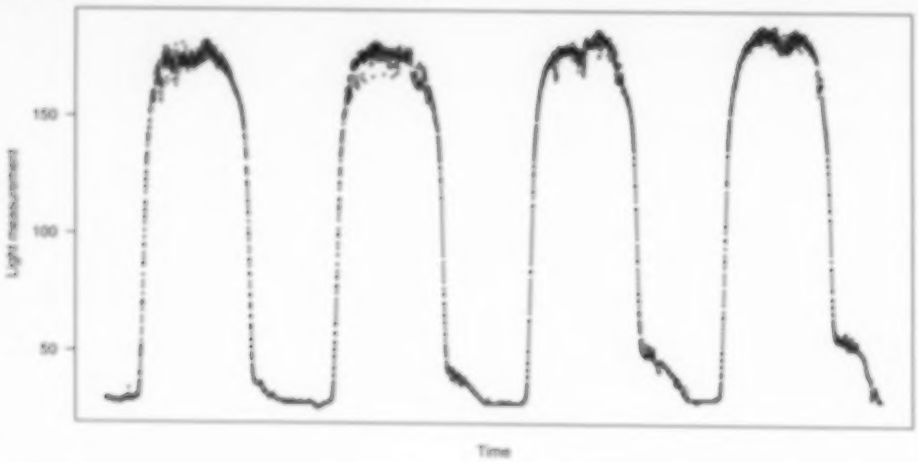
Received 19 December 2005. Accepted 27 June 2006. Published on the NRC Research Press Web site at http://cjfas.nrc.ca/ on 1 August 2007.
J19554

A. Nielsen[1] and J. R. Sibert. Pelagic Fisheries Research Program, Joint Institute for Marine and Atmospheric Research, University of Hawai'i at Mānoa, 1000 Pope Road, MSB 312, Honolulu, HI 96822, USA.

[1]Corresponding author (e-mail: anders.nielsen@hawaii.edu).

 doi: 10.1139/f07-064 © 2007 NRC Canada

Fig. 1. Four days of raw light measurements from an archival tag.



time of threshold crossing near sunset and delay it near sunrise, so the latitude determined by the difference will be doubly wrong. Around an equinox, where day lengths are very similar at all latitudes, even small errors in the crossing times will result in huge errors in the estimated latitudes. The longitudinal estimates are generally more accurate, as they are computed from the midpoint of the two daily crossing times. If one crossing time is delayed and the other is equally advanced, the longitudinal estimate is still correct.

A possible improvement is presented by Ekstrom 2004, where it is suggested to fit a template function to narrow-band blue light in the twilight period. In this period, light in these wavelengths may be less sensitive to weather conditions.

Many variants of threshold algorithms are in use, but the details of some algorithms are proprietary, which makes it impossible to present the algorithms here and to thoroughly study them. What can be studied is their performance. Often, tagged marine animals are mistakenly placed hundreds of kilometres from their actual location, which can easily be on dry land (e.g. Welch and Eveson 1999; Musyl et al. 2001; and Sibert et al. 2003). Close to equinoxes, these algorithms frequently fail completely for the latitude coordinate and produce geolocations that are up to 40° of latitude from the true locations (Musyl et al. 2001).

Intuitively the solution is simple. Each position should not be estimated in a knowledge void. The first position is the known release position. The second position must be expected to be in the neighborhood of the first, and so on. A statistically formal way to express this is a state–space model (Harvey 1990). A state–space model works by assuming an underlying movement model, and estimates the track that best matches both the assumed model and the available data. Even a simplistic movement model can be helpful. Each point on the estimated track is a weighted average of the prediction from the assumed movement model and the estimate from the daily light data. The weighting is done according to prediction precision of the assumed movement model and the precision of the data.

The first state–space model for light-based geolocation (Sibert et al. 2003) used the raw light-based daily geolocations as the only data and assumed a simple random walk as the movement model. This model greatly improved the estimated tracks and became widely used (e.g. Musyl et al. 2003, Wilson et al. 2005, and Wilson et al. 2006). The model was extended to also use sea surface temperature (Nielsen et al. 2006), which further improved the estimated track, especially the latitude coordinate in those periods around each equinox where the light-based information is most uncertain.

Using the two step approach, of first estimating raw geolocations by threshold algorithms and then using these estimates as data in a state–space model, is not optimal. The threshold algorithms are not helped by information from previous positions, as each raw geolocation is processed independently. This leads to great uncertainties, especially around equinoxes. Furthermore, the two step approach prevents the state–space model from getting access to all the relevant light data, as the light data are reduced to a (possibly poor) geolocation before it is used in the state–space model. One obvious consequence of this is that any information about the uncertainty of each raw geolocation that might have been in the light data is lost, and the state–space model must re-estimate this information.

This paper presents a coherent state–space model that estimates a track of geolocations directly from light measurements in the tag. All problems associated with threshold algorithms are completely avoided, as they are not a part of the model. This allows all parts of the model movement, covariance structure, and even the functional relation between solar altitude and light measurements to be estimated within the model. The model gives a sound statistical foundation for estimation of the most probable track, and for quantifying its precision.

Materials and methods

The raw record of an archival tag is a series of measurements taken at times $t_1, t_2, \ldots, t_M$. Typically these times are equidistant, but values can be missing. At these times, three variables are measured: depth, temperature, and light intensity. In the following, only time and light intensity (t, I) are considered.

Focus on solar event periods

From the raw light record (example in Fig. 1), periods of daytime and nighttime can easily be identified and in between are the periods surrounding sunrise and sunset. Light intensity

© 2007 NRC Canada

is only useful for geolocating if a change in position translates to a change in light intensity. At midnight, even a fairly big change in position (say 5°) will not change the light intensity noticeably — it will still be dark. Similarly at noon, the change in solar altitude resulting from changing position a few degrees will change the light intensity slightly, but local weather conditions, cloud cover, and other uncontrollable features dominates the measured light intensities completely. The periods around solar events (sunrises and sunsets) are the relevant parts for geolocation.

An automatic procedure to select the critical periods around each solar event can be set up in many ways. It is likely not important which procedure is chosen, as long as the resulting intervals include the informative parts of data. The following describes the procedure used in this paper.

First, the raw light record is partitioned into 24 h periods. Within each of these 24 h periods, two observation times are selected such that the mean of the light measurements between the two times ($\widetilde{t}_i$ and $\widetilde{t}_{i+1}$) and the mean of the remaining light measurements times differ the most. In other words:

$$(1)\qquad \widetilde{t}_i, \widetilde{t}_{i+1} = \underset{a,b}{\operatorname{argmax}} \left|\operatorname{mean}(\{l_j : a < t_j \le b\}) - \operatorname{mean}(\{l_j : t_j \le a \vee b < t_j\})\right|$$

The times $\widetilde{t}_1, \widetilde{t}_2, \ldots, \widetilde{t}_{2N}$, where N is the number of days in the light record, are used to select the data intervals to be used in the geolocation model. The interval around each solar event i is chosen asymmetrically to be from 14.4 min (1% of 24 h) towards the nighttime from $\widetilde{t}_i$ to 72 min (5% of 24 h) towards the daytime from $\widetilde{t}_i$. The interval is chosen asymmetrically, because it is not desirable to include points from the dark nighttime period or close to the nighttime period, as moonlight can bias the light measurements in this period. Choosing the length of these intervals (here 86.4 min) is a trade-off. Longer intervals include more data in the analysis, but also increase computation times substantially. Furthermore, the model assumes that horizontal movement within each interval is negligible, and shorter intervals make this assumption more reasonable. Threshold algorithms assume that movement between two solar events is negligible, which is clearly less reasonable.

Observation times within the ith interval are denoted $\tau^{(i)} = (\tau_1^{(i)}, \ldots, \tau_{n_i}^{(i)})$, and the corresponding light observations are denoted $l^{(i)} = (l_1^{(i)}, \ldots, l_{n_i}^{(i)})$. Finally, the average observation time within the ith interval is denoted $\bar{t}_i$. These average times $\bar{t}_1 \le \cdots \le \bar{t}_{2N}$ will be the times where geolocations are computed by the following model (illustrated in Fig. 2).

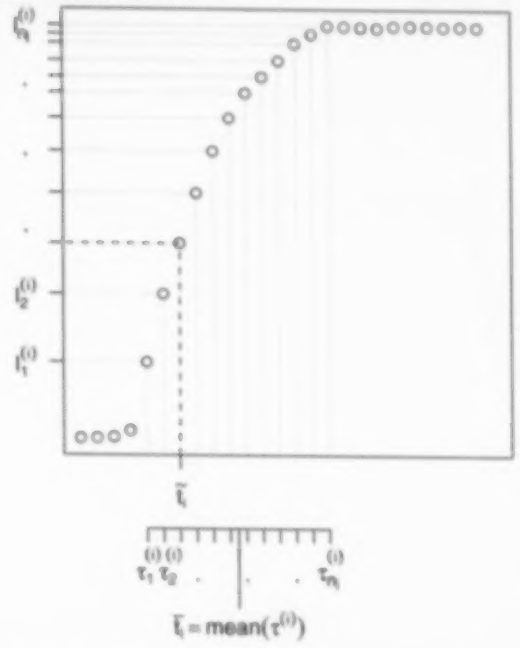
Model

The model for the observed light measurements along an animal track is a state–space model, where the transition equation describes the movements along the sphere, and the measurement equation describes the predicted light measurements at any given position. For a fixed set of model parameters, the model reconstructs the track where the movements between positions best matches the transition equation, and where the positions yields the best predictions of the actual light measurements.

For the transition equation, a random walk model is assumed:

$$(2)\qquad \alpha_i = \alpha_{i-1} + c_i + \eta_i, \qquad i = 1, \ldots, 2N$$

Fig. 2. Illustration of the relationship among $\widetilde{t}_i$, $\tau^{(i)}$, and $\bar{t}_i$ around one solar event.



Here, α_i is a two-dimensional vector containing the coordinates $(\alpha_{i,1}, \alpha_{i,2})$ in nautical miles (1 nautical mile = 1.853 km) along the sphere from a translated origin at time $\bar{t}_i$. c_i is the drift vector describing the deterministic part of the movement, η_i is the noise vector describing the random part of the movement, and N is the number of days in the track. The deterministic part of the movement is assumed to be proportional to time $c_i = (u\Delta\bar{t}_i, v\Delta\bar{t}_i)'$. The random part is assumed to be serially uncorrelated and follow a two-dimensional Gaussian distribution with mean vector 0 and covariance matrix $Q_i = 2D\Delta\bar{t}_i I_{2\times 2}$. Here, D is a model parameter expressing the diffusive movement component and $I_{2\times 2}$ is the two-dimensional identity matrix.

The measurement equation of the state–space model describes the expected light measurements at the position α_i at times $\tau^{(i)}$. The calculation of the expected light measurements can be partitioned into three steps. (*i*) Position α_i is transformed into degrees of longitude and latitude by a function z as described in (Sibert et al. 2003). (*ii*) Solar altitudes θ at the position $z(\alpha_i)$ at times $\tau^{(i)}$ are calculated from standard astronomical algorithms (Meeus 1998). (*iii*) Expected light measurements are calculated from the solar altitudes. The function φ describing this relationship is not a known function, and hence it must be estimated within the model.

The function φ is represented in the model by a cubic spline function interpolating the points $(\widetilde{\theta}_1, \widetilde{\varphi}_1), \ldots, (\widetilde{\theta}_{n_\varphi}, \widetilde{\varphi}_{n_\varphi})$, where $\widetilde{\theta} \le \cdots \le \widetilde{\theta}_{n_\varphi}$ are chosen equidistant over the angles in the data. Notice here that the selection of data intervals around solar events substantially narrows the angle interval where this function is needed. The $\widetilde{\varphi}$s are model parameters. To stabilize the optimization and to reflect our knowledge about the relation between solar altitude and light intensities the $\widetilde{\varphi}$s are restricted to be increasing. Combining steps *i*–*iii* leads to the following

© 2007 NRC Canada

measurement equation:

$$(3) \qquad l^{(i)} = \varphi_{\tilde{\theta}}(\theta(z(\alpha_i), \tau^{(i)})) + \varepsilon_i, \qquad i = 1, \ldots, 2N$$

For easier future reference, the entire nonlinear mapping from position to expected light measurement is denoted Λ, which reduces the measurement equation to

$$(4) \qquad l^{(i)} = \Lambda(\alpha_i, \tau^{(i)}) + \varepsilon_i, \qquad i = 1, \ldots, 2N$$

The measurement error ε_i is assumed to follow a Gaussian distribution with mean vector 0 and covariance matrix $\Sigma^{(i)}$, where

$$(5) \qquad \Sigma^{(i)}_{j,k} = \begin{cases} \sigma_1^2 + \sigma_2^2 + \sigma_3^2, & \text{if } j = k \\ \sigma_1^2 + \sigma_2^2 \exp(-|\tau_j^{(i)} - \tau_k^{(i)}|/\rho), & \text{if } j \neq k \end{cases}$$

and $j, k = 1, \ldots, n_i$. Here, $\sigma_1, \sigma_2, \sigma_3$, and ρ are model parameters. Each element $\Sigma^{(i)}_{j,k}$ in this covariance matrix describes the covariance between two light measurements $l_j^{(i)}$ and $l_k^{(i)}$ taken around the same solar event.

The covariance structure in eq. 5 reflects the intuition that two light measurements taken near the same solar event are more similar (correlated) than two taken at separate solar events (σ_1). If visibility is low one morning, all measurements will be lowered. The covariance structure furthermore allows the correlation between two light measurements near the same solar event to decrease as the time between them increases (σ_2). Finally, the covariance structure includes a term describing the independent measurement errors (σ_3).

Now the state–space model is completely defined. All parameters of this model are

$$(6) \qquad \vartheta = (u, v, D, \tilde{\varphi}_1, \ldots, \tilde{\varphi}_{n_g}, \sigma_1, \sigma_2, \sigma_3, \rho)$$

The following describes how the model is optimized with respect to these model parameters and how the model is used to reconstruct the most probable track.

Estimation via the unscented Kalman filter

The basic Kalman filter (Harvey 1990) assumes that both the transition equation and the measurement equation of the state–space model are linear. The extended Kalman filter (Harvey 1990) can handle slight nonlinearities by local first order Taylor approximations of the nonlinear functions in the model. The unscented Kalman filter (Julier et al. 2000) is a more recent sequential estimation technique. Instead of approximating the nonlinear functions, the transformed probability distributions are approximated directly. This approach gives a simpler implementation, not requiring derivatives of the equations in the state–space model; it also allows for higher accuracy — at least corresponding to a second-order Taylor approximation (Julier et al. 2000).

The unscented transformation

Assume that x is a d-dimensional random variable following a Gaussian distribution with mean vector μ_x and covariance matrix Σ_x. Consider now the distribution of an arbitrary, nonlinear transformation $y = f(x)$

The first-order Taylor approximation, which is used in the extended Kalman filter, states that

$$(7) \qquad y \overset{\text{approx}}{\sim} N\left(f(\mu_x), \left(\left.\frac{df}{dx}\right|_{x=\mu_x}\right)\Sigma_x\left(\left.\frac{df}{dx}\right|_{x=\mu_x}\right)'\right)$$

The unscented transformation is a different approximation of the distribution of y. The distribution of x is represented by a set $\{x\}$ of $2d$ cleverly selected points:

$$(8) \qquad \{x\} \overset{\text{def}}{=} \mathrm{SP}(\mu_x, \Sigma_x) \overset{\text{def}}{=} \left\{ x \,\middle|\, x = \mu_x \pm \left(\sqrt{d\Sigma_x}\right)_i, \quad i = 1, \ldots, d \right\}$$

where $(\sqrt{d\Sigma_x})_i$ denotes the ith column of the Cholesky decomposition of the matrix $d\Sigma_x$.

These so-called sigma points are a simple transformation of the mean and covariance of the distributions they represent, and the notation SP in eq. 8 is the shorthand notation for this transformation.

Each of these sigma points are now transformed by the function f to their corresponding y points:

$$(9) \qquad y^{(j)} = f(x^{(j)}), \qquad j = 1, \ldots, 2d$$

This set of y points representing the transformed distribution is denoted $\{y\}$, and a shorthand for this transformation of x points into y points, which is used in the following, is defined as $\{y\} = f(\{x\})$.

The mean vector (eq. 10) and covariance matrix (eq. 11) of the transformed points are now computed as

$$(10) \qquad \hat{\mu}_y = \mathrm{mean}(\{y\}) = \frac{1}{2d}\sum_{j=1}^{2d} y^{(j)}$$

$$(11) \qquad \hat{\Sigma}_y = \mathrm{cov}(\{y\}) = \frac{1}{2d}\sum_{j=1}^{2d} (y^{(j)} - \hat{\mu}_y)(y^{(j)} - \hat{\mu}_y)'$$

The approximate distribution of y given by the unscented transformation is the Gaussian distribution with mean vector $\hat{\mu}_y$ and covariance matrix $\hat{\Sigma}_y$. With the notation introduced here, the entire unscented transformation can be summarized in

$$(12) \qquad y \overset{\text{approx}}{\sim} N\big(\mathrm{mean}(f(\mathrm{SP}(\mu_x, \Sigma_x))), \mathrm{cov}(f(\mathrm{SP}(\mu_x, \Sigma_x)))\big)$$

Notice when comparing eq. 12 with eq. 7 that the approximation in the unscented transformation is simpler to compute, as it does not require derivatives of f. All that is needed is to evaluate the nonlinear function in a few carefully selected points.

The unscented transformation is presented here in its simplest form. Other versions include optional scaling of the sigma points. This is introduced to keep the sigma-points closer together when the dimension of x is high. In this application $d = 2$, so these complications are avoided.

The unscented Kalman filter

The unscented Kalman filter is very similar to the extended Kalman filter, except that the unscented transformation is used instead of local Taylor approximations. The following describes the exact unscented Kalman filter equations for this model.

Step 0: The filter is started by assuming that the first position $\hat{\alpha}_0$ in the track is known without error (its covariance is $P_0 = 0_{2\times 2}$). This is a reasonable assumption, as the first position is the release position. The initial set of sigma points is computed:

$$(13) \qquad \{\alpha_0\} = \mathrm{SP}(\hat{\alpha}_0, 0_{2\times 2}), \text{ and set } i = 1$$

© 2007 NRC Canada

Step 1: From the current set of sigma points and the transition equation (eq. 2), the prediction of the next position is calculated, along with its covariance and its sigma point representation:

$$\widehat{\alpha}_{i|i-1} = \text{mean}(\{\alpha_{i-1}\} + c_i) \quad (14)$$

$$P_{i|i-1} = \text{cov}(\{\alpha_{i-1}\} + c_i) + Q_i \quad (15)$$

$$\{\alpha_{i|i-1}\} = \text{SP}(\widehat{\alpha}_{i|i-1}, P_{i|i-1}) \quad (16)$$

Step 2: The predicted light level vectors (eq. 17) corresponding to the sigma points are calculated via eq. 4, along with their mean vector, covariance, cross covariance, and a matrix known as the Kalman gain matrix, which is used when updating the predicted position with the actually observed light vector:

$$\{l^{(i)}\} = \Lambda(\{\alpha_{i|i-1}\}, \tau^{(i)}) \quad (17)$$

$$\widehat{l}^{(i)} = \text{mean}(\{l^{(i)}\}) \quad (18)$$

$$F_i = \text{cov}(\{l^{(i)}\}) + \Sigma^{(i)} \quad (19)$$

$$G_i = \text{cov}(\{\alpha_{i|i-1}\}, \{l^{(i)}\}) \quad (20)$$

$$K_i = G_i F_i^{-1} \quad (21)$$

Step 3: Finally, the difference between the predicted light vector and the observed is calculated, the Kalman update equations are applied, and an updated set of sigma points are calculated:

$$w_i = l^{(i)} - \widehat{l}^{(i)} \quad (22)$$

$$\widehat{\alpha}_i = \widehat{\alpha}_{i|i-1} + K_i w_i \quad (23)$$

$$P_i = P_{i|i-1} - K_i F_i K_i' \quad (24)$$

$$\{\alpha_i\} = \text{SP}(\widehat{\alpha}_i, P_i) \quad (25)$$

Step 4: If $i < 2N$, increase i by one and go to step 1.

Negative log-likelihood function

The maximum likelihood principle is used to find the estimates of all model parameters (eq. 6). With the definition that $w_0 = l^{(0)} - \Lambda(\widehat{\alpha}_0, \tau^{(0)})$ and $F_0 = \Sigma^{(0)}$, the negative log-likelihood is given by

$$\ell(\vartheta) = -\sum_{i=0}^{2N} \log \phi(w_i, 0, F_i) \quad (26)$$

Here, $\phi(w, 0, F)$ is the density of the Gaussian distribution with mean vector 0 and covariance matrix F evaluated in w.

Last point known without error

Knowing the last position in the track corresponds to extending the last observational vector $l^{(2N)}$ with the known two-dimensional position to $l^{(2N)} = (l^{(2N)}, \alpha_{2N})'$ (here expressed in nautical miles coordinates). The model extension to accommodate this extra information is straightforward. Λ is extended to also return the predicted position, and the observational covariance matrix $\Sigma^{(2N)}$ is extended by two zero-rows and two zero-columns to express that the position is known without error. Everything else remains unchanged.

Most probable track

The unscented Kalman filter and the maximum likelihood principle supply estimates of the model parameters and the predicted track. A point on the predicted track at any given time point is calculated using all observations available at that time, that is, $\widehat{\alpha}_i \approx E(\alpha_i | l^{(1)}, \ldots, l^{(i)})$. Once the entire set of light observations are available, it is possible to obtain estimates with smaller variance. What is here denoted "the most probable track" is calculated using all observations, after the parameters have been estimated. A point on the most probable track is $\widehat{\alpha}_{i|2N} \approx E(\alpha_i | l^{(1)}, \ldots, l^{(2N)})$.

The actual computation of the most probable track is done in a single backwards sweep of the predicted track and its covariance. The last point of the most probable track is identical to the last point of the predicted track, as all observations were available to the predicted track at the final point. The last point of the most probable track is $\widehat{\alpha}_{2N|2N} = \widehat{\alpha}_{2N}$, and its covariance is $P_{2N|2N} = P_{2N}$. The following equations are used to recursively compute the previous points of the most probable track:

$$\widehat{\alpha}_{i|2N} = \widehat{\alpha}_i + P_i^*(\widehat{\alpha}_{i+1|2N} - \widehat{\alpha}_i - c_{i-1}) \quad (27)$$

$$P_{i|2N} = P_i + P_i^*(P_{i+1|2N} - P_{i+1|i})P_i^{*\prime} \quad (28)$$

$$\text{where} \quad P_i^* = P_i P_{i+1|i}^{-1}$$

In general, it is not possible to use these simple formulas with the unscented Kalman filter, but the transition equation (eq. 2) of this model is linear, so here it is a valid approximation. A general approach is described in Wan and van der Merwe (2001). In textbooks on the Kalman filter (Harvey 1990), this technique is known as smoothing, as the resulting track most often is smoother than the predicted track.

Simulation of realistic data

Simulating artificial, but realistic, data from the actual model, is an important tool to investigate most important properties of the model. The advantage of using simulated data is that they are cheap, the true tracks and model parameters are always known, and they can be tailored to focus on specific aspects of the model.

From a set of known model parameters ϑ and a known first position $\alpha_0 = z^{-1}(y_0)$, the following steps are used to simulate data.

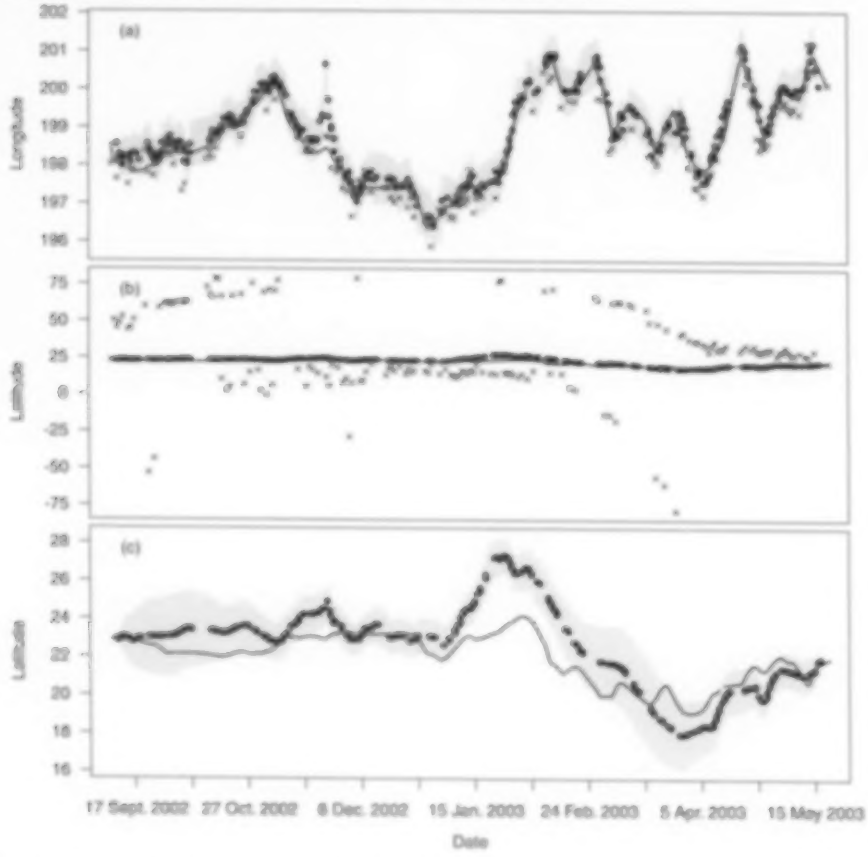
(*i*) The underlying true movements are simulated forward in time from the first position via the movement parameters (u, v, D) and the transition equation (eq. 2). The times used $(t_1, t_2, \ldots, t_M)$ should mimic light sampling rates from a real archival tag. Depending on tag type, they could be 20 s or 5 min apart.

(*ii*) At each point in the true simulated track, the predicted light intensity is calculated via the function Λ in eq. 4 and the parameters $\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_{n_\varphi}$ describing the function φ relating the solar altitude to the tag-measured light intensity. Real estimates of model parameters from an actual tag are used here to make the simulated data as realistic as possible.

(*iii*) The uncorrelated part of the observation noise, $\varepsilon_i \sim N(0, \sigma_i^2), t_1, \ldots, t_M$, is now added to each of these light predictions.

© 2007 NRC Canada

Fig. 3. Points on the most probable track from the light-based state–space model (circles), along with raw geolocations from the tag manufacturer's threshold algorithm (crosses) and the true track from the satellite-based method (solid line). The shaded areas are 95% confidence regions estimated by the light-based state–space model. Data indicate longitude coordinate (*a*); latitude coordinate including raw geolocations (*b*); and latitude zoomed to scale of most probable and true track (*c*).



(*iv*) The scanning algorithm, previously described, is applied to select the solar event times and the surrounding times and simulated light intensities.

(*v*) Finally, the remaining part of the observation noise is added. To the simulated light intensities surrounding the ith solar event is added a vector $e^{(i)} \sim N(0, S^{(i)}), i = 1, \ldots, 2N$, where $S^{(i)}$ is equal to $\Sigma^{(i)}$ described in eq. 5, but with $\sigma_\lambda^2 = 0$.

The resulting data are a set of time vectors and matching simulated light observation vectors $(t^{(i)}, l^{(i)})_{i=1,\ldots,2N}$. These artificial observations will follow the model, except for one minor detail. The model assumes no movement within the time intervals surrounding each solar event, but the simulation movement continues in these intervals. This is done to better mimic real data, and to investigate the consequences of this model assumption.

Results

A Wildlife Computers pop-up archival transmitting tag (version 2) was deployed on a drifter buoy near Hawaii in September 2002 and set to pop up after 9 months. The software from the manufacturer was used to estimate threshold-based geolocations. The resulting latitude reconstruction was extremely inaccurate (Fig. 3). The data from this type of tag consist of ~ 12 light measurements around each solar event, taken ~ 8 min apart. These raw light measurements were used in the described light-based state–space model, without first applying the scanning described. The estimated most probable track, especially the latitude coordinate, was greatly improved (Fig. 3). The estimated 95% confidence region around the most probable track is not perfect, but it does seem to have approximately the correct size. The tag manufacturer's algorithm for selecting the light measurements to be included in the data does not seem as robust as the one described in this paper, so even better results can be expected if the observation window was selected using that instead.

An important pattern is clearly identified in the standard deviation of the most probable track (Fig. 4). The standard deviation of the longitude is fairly constant throughout the year, but the latitude standard deviation is increasing (doubling) around the two equinoxes. This pattern is not assumed in the model, as was necessary in Sibert et al. (2003) and Nielsen et al. (2006), but is a simple consequence of propagating the data uncertainties

© 2007 NRC Canada

Fig. 4. Estimated standard deviation (SD) of the geolocations from the light-based state–space model. Vertical broken lines indicate time of equinox. Data indicate longitude coordinate (*a*) and latitude coordinate (*b*).

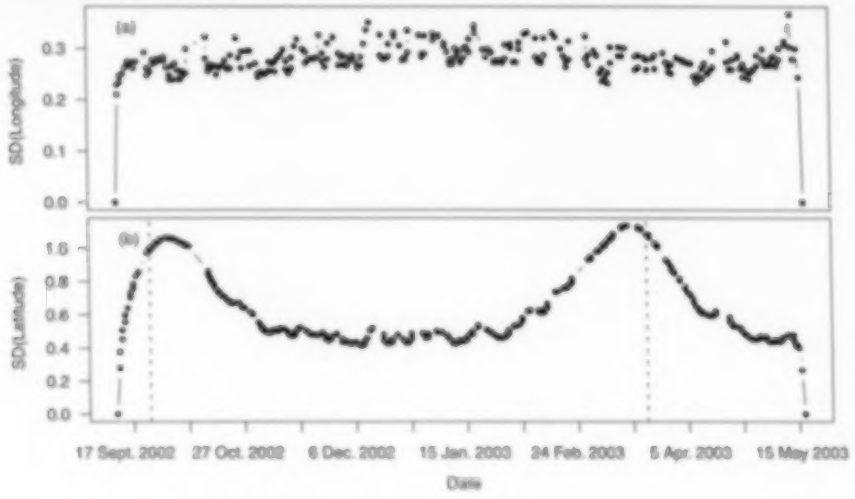
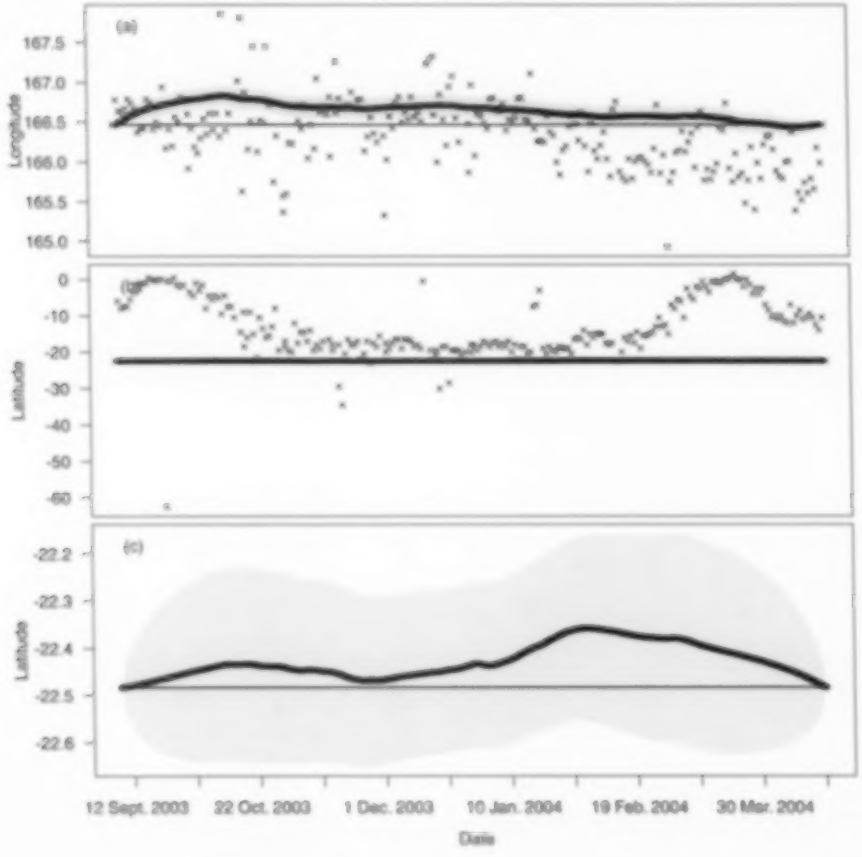
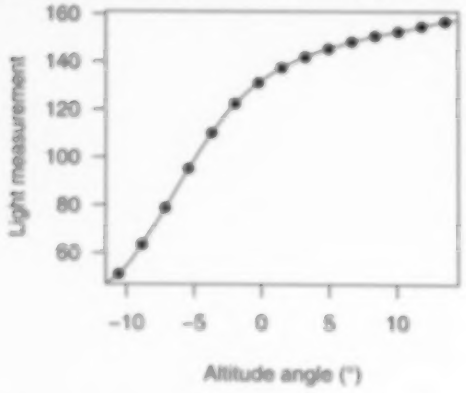


Fig. 5. Points on the most probable track from the light-based state–space model (circles here are so close that they appear as a thick solid line), along with raw geolocations from the tag manufacturer's threshold algorithm (crosses) and the true track from satellite-based method (solid line). The shaded areas are 95% confidence regions estimated by the light-based state–space model. Data indicate longitude coordinate (*a*); latitude coordinate including raw geolocations (*b*); and latitude zoomed to scale of most probable and true track (*c*).



© 2007 NRC Canada

Fig. 6. The estimated function φ describing the relationship between solar altitude angle and the expected light measurement in the tag. The circles are the estimated support points corresponding to the model parameters $\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_{n_\varphi}$.



through geolocation process.

A Wildlife Computers archival tag (Mk9, version 1.03) was attached to a mooring near New Caledonia in September 2003, and recorded data for close to 8 months. The software from the manufacturer was used to estimate threshold-based geolocations. The reconstructed latitude was extremely inaccurate (Fig. 5). The raw data contains a light reading every minute. The most probable track from the light-based state–space model is significantly biased for the longitude coordinate in the beginning of the track, whereas the threshold-based longitudes are similarly biased near the end of the track. This bias is likely a result of the internal clock in the tag being slightly biased or drifting. The latitude coordinate of the most probable track is within 0.2° of the true position at all times, and the latitude confidence region covers the true position. Even including the longitude bias, the most probable track is within 0.5° of the true position at all times.

The function φ describing the relationship between the solar altitude and the expected light measurement is an important part of the model. The function is estimated within the model for the interval covering the selected light measurements. The function appears to reliably estimated, as the estimated support points are in close agreement with a monotonic increasing curve (Fig. 6). The function φ estimated from this mooring study is used as the basis for the following simulations.

To test the model with realistic movement parameters, two scenarios were simulated: first, with a moderate diffusion rate of $D = 100$ Nm^2·day^{-1}, which is in the range of sea turtle movements (Y. Swimmer, NOAA Fisheries, Pacific Island Fisheries Science Center, 2570 Dole Street, Honolulu, HI 96822, Yonat.Swimmer@noaa.gov, personal communication); second, with a high diffusion rate of $D = 500$ Nm^2·day^{-1}, which is in the range of bigeye tuna (*Thunnus obesus*) movement (Sibert et al. 2003). Three tracks, at three different release latitudes, were simulated for each scenario. The actual estimated values from the mooring data were used for the functional relationship between the solar altitude and light (Fig. 6) and for the parameters determining the variance structure ($\sigma_1^2 = 83$, $\sigma_2^2 = 9.9$, $\sigma_3^2 = 2.6$, and $\rho = 3.2\%$). This makes the conclusions based on these simulations regarding precision of the most probable track as realistic as possible.

Table 1. True parameter values, means, and standard deviations (SD) of estimates from 100 simulated tracks.

Parameter name	True value	Mean of estimates	SD of estimates
u	0.000	−0.138	1.366
v	0.000	0.123	1.310
D	300.000	295.021	34.916
ρ	0.032	0.034	0.006
σ_1^2	83.000	81.829	5.041
σ_2^2	9.900	10.424	1.281
σ_3^2	2.600	2.606	0.075
φ_{-5}	98.843	98.793	0.550
φ_0	132.255	132.242	0.407
φ_5	145.571	145.487	0.353

The moderately moving tracks were all estimated accurately (Fig. 7). The longitude coordinate is within 0.5° of the true track, and the latitude coordinate is within 1° of the true track. The estimated 95% confidence region appears to have the correct coverage. The standard deviation of the most probable latitude show an interesting pattern. Around each equinox, the standard deviation increases, but not in the same way for all three tracks. The track near the equator (panels *a* and *d*) has the highest latitude standard deviation exactly at the equinox, but for the northern track (panels *b* and *e*) and the southern track (panels *c* and *f*), the time of the highest latitude standard deviation is shifted towards the winter side. Further, it is noted that the standard deviation is higher for the track near equator, where annual variation in day length is least.

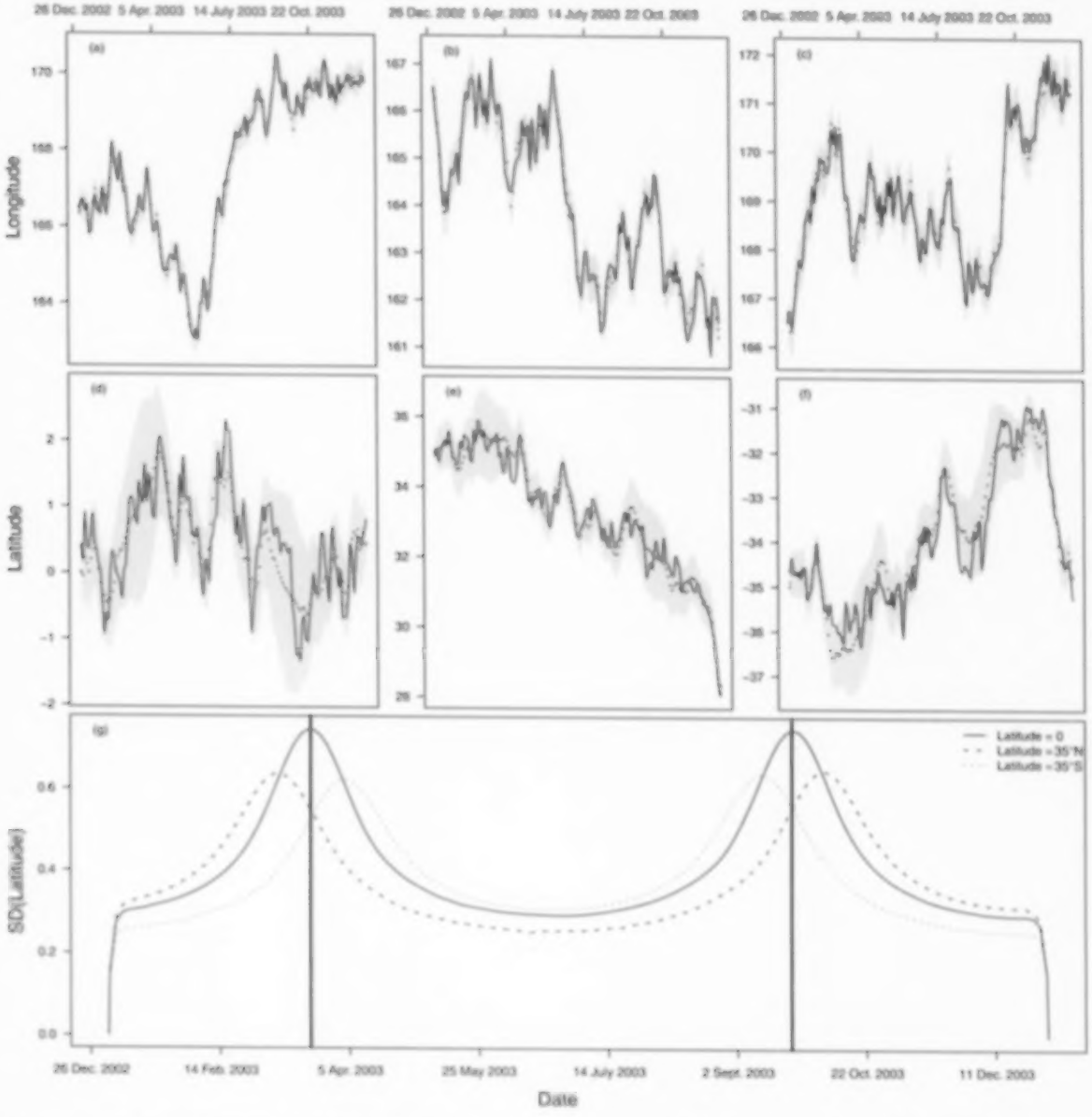
The fast-moving tracks were also estimated accurately (Fig. 8). The longitude coordinate is within a degree of the true track, and the latitude coordinate is within 2° of the true track, but most often closer. The estimated 95% confidence region appears to have the correct coverage. The standard deviation of the most probable latitude standard deviation shows the same pattern as for the moderately moving tracks, except that the standard deviation is higher for the fast-moving tracks. This is expected, as larger diffusion D leads to larger uncertainty in the prediction from the underlying movement model.

To verify that all model parameters in this model are identifiable, 100 tracks were simulated from known parameters, and the light-based state–space model was applied to each track. The parameters representing the relationship between solar altitude and light $\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_{n_\varphi}$ are chosen dynamically to cover the range (of angles) of the selected light measurements, and hence they may represent the light value at slightly different angles for each simulated track. To be able to verify the estimated relationship between between solar altitude and light three fixed angles are selected: −5°, 0°, and 5°. For each simulated track, the estimated light levels (φ_{-5}, φ_0, and φ_5) at these three angles are evaluated.

The simulation study shows that all model parameters are identifiable. Including the relationship between solar altitude and light (Table 1), the movement parameters (u, v, and D) show no significant bias. The relationship between solar altitude and light is estimated unbiased at −5° and 0°, but these 100 simulations indicate a slightly biased estimate (of 0.084 light units) at 5°, but this is likely a coincidence occurring because of the relative low number of simulations.

© 2007 NRC Canada

Fig. 7. Longitude (*a*, *b*, and *c*) and latitude (*d*, *e*, and *f*) of three moderately moving simulated tracks: released at equator (*a* and *d*), released at 35°N (*b* and *e*), and released at 35°S (*c* and *f*). Each frame shows the true simulated track (solid line), the most probable track from the light-based state–space model (dots), and the 95% point-wise confidence region (shaded area). Finally, the estimated standard deviation (SD) of the most probable latitude is shown (*g*).

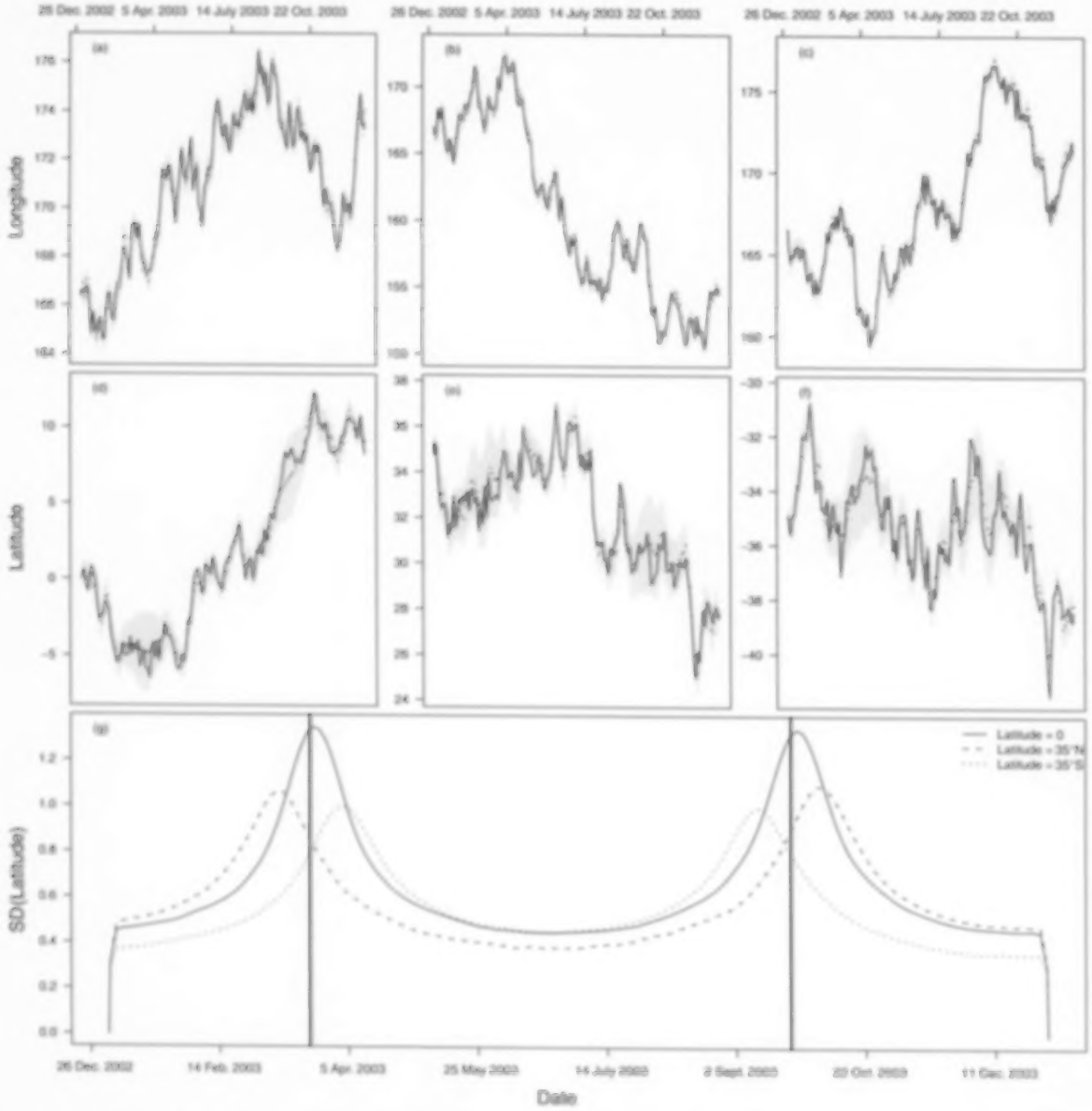


Of the estimated parameters in the variance structure (σ_1, σ_2, σ_3, and ρ), three are slightly biased. The parameter describing the variation in visibility between different solar events (σ_1) is negatively biased, but the parameters describing the correlation between close observations within the same solar event (σ_2 and ρ) are positively biased. The estimates are trading a small part of the overall correlation between light measurements from the same solar event to higher correlation between neighboring points. This is a consequence of the model assumption of no movement within the observational window around each solar event. The simulated tracks, and likely also real tracks, continue to move around solar event times, which gives this higher correlation between neighboring points. The actual value of these parameters are rarely used, so this small bias is unproblematic.

In all previous simulations (Figs. 7 and 8; Table 1), the sampling rate has been set to 300 measurements per day, which is one every 4.8 min. The observation window refers to the interval around each solar event where the light measurements are

© 2007 NRC Canada

Fig. 8. Longitude (*a*, *b*, and *c*) and latitude (*d*, *e*, and *f*) of three fast-moving simulated tracks: released at the equator (*a* and *d*), released at 35°N (*b* and *e*), and released at 35°S (*c* and *f*). Each frame show the true simulated track (solid line), the most probable track from the light-based state–space model (dots), and the 95% point-wise confidence region (shaded area). Finally, the estimated standard deviation (SD) of the most probable latitude is shown (*g*).



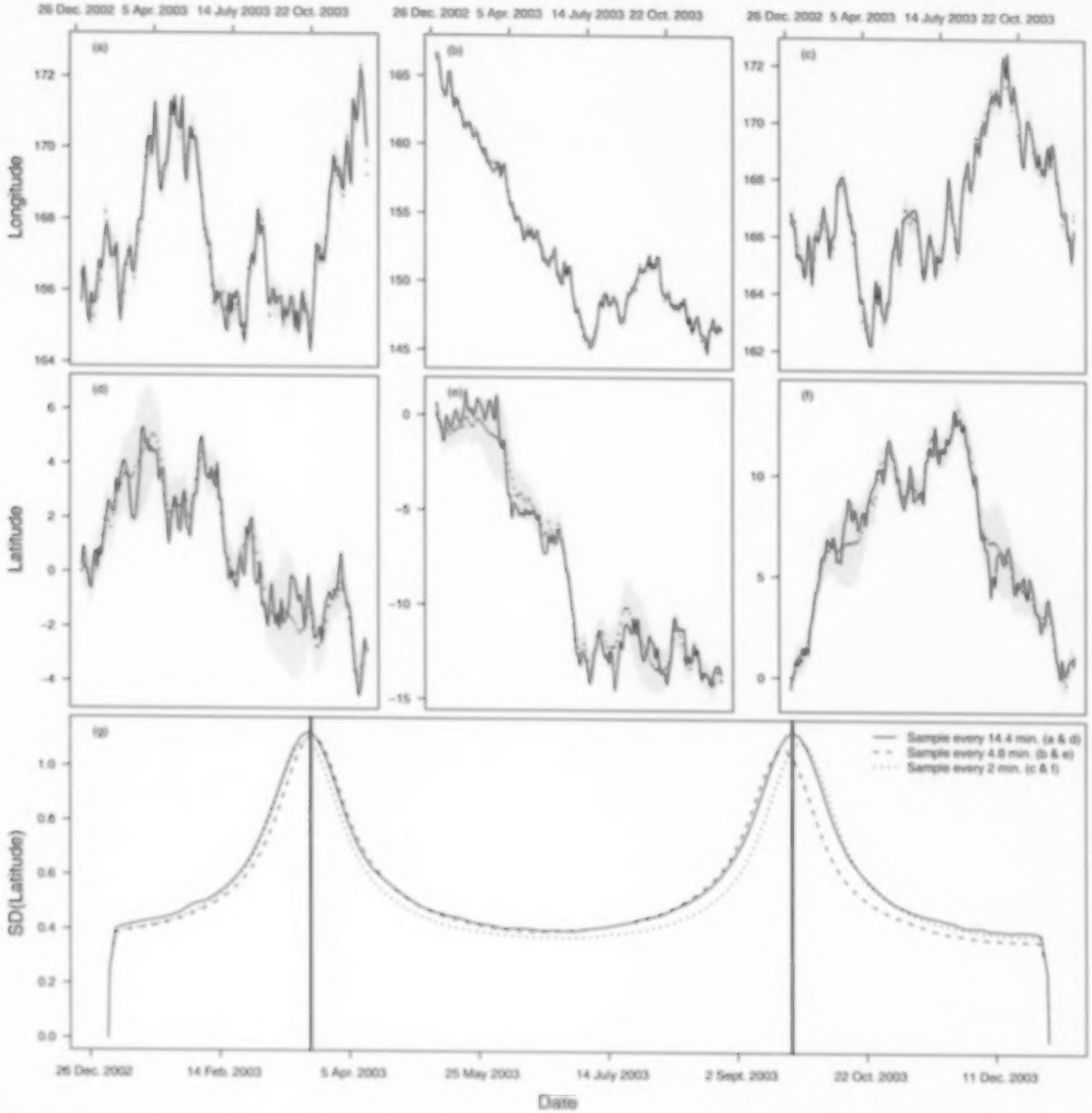
actually used by the model. The observation window, consisting of all observations from the solar event time and 72 min towards the daytime, and all observations from the solar event and 14.2 min towards the nighttime has previously been used.

The sampling rate was varied from every 2 min to every 14.2 min. In all cases, the model was able to follow the true track (Fig. 9). The standard deviation of the most probable latitude was higher at the sparse sampling rates, but the seasonal differences are much more dominant (Fig. 9).

The observation window was tested by using fewer of the observations on the day side of the solar event, or fewer of the observations on the night side of the solar event, or both. In all cases, the model was still able to follow the true track (Fig. 10). The standard deviation of the most probable latitude was higher in the case where only 18 min towards the day and 3.6 min towards the night was used (one-quarter of the original window), but all other tested window sizes gave surprisingly similar latitude variances.

© 2007 NRC Canada

Fig. 9. Longitude (*a*, *b*, and *c*) and latitude (*d*, *e*, and *f*) of three simulated tracks: light sampling every 2 min (*a* and *d*), every 4.8 min (*b* and *e*), and every 14.2 min (*c* and *f*). Each frame shows the true simulated track (solid line), the most probable track from the light-based state–space model (dots), and the 95% point-wise confidence region (shaded area). Finally, the estimated standard deviation (SD) of the most probable latitude is shown (*g*).



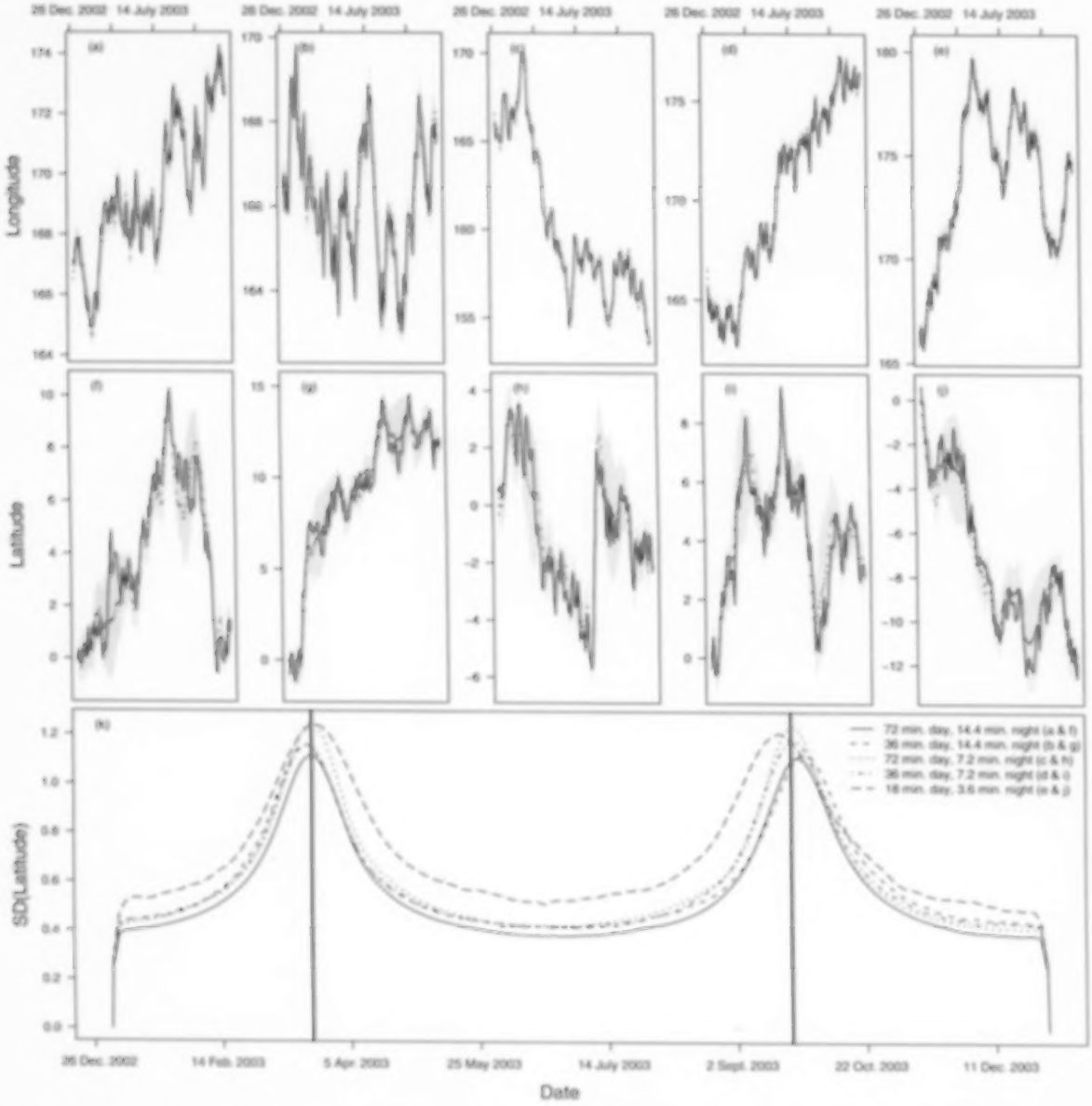
Discussion

The model presented in this paper is completely different from previous algorithms used in this field. It is unique in combining all steps from the raw light measurements to the estimated most probable track in one coherent model. It cannot be classified as a threshold model, as it does not reduce the light measurements at each solar event to a time where a certain solar altitude is assumed known. It cannot be classified as a template-matching method (as described in Ekstrom 2004) either, as it does not assume to know the relationship between solar altitude and light. Instead it estimates this relationship within the model.

The model presented here is also unique in offering a way to deal with correlated light measurements. It is a fact that light measurements are serially correlated. Assuming they are independent will lead to false confidence statements. The value of each light measurement is influenced by many factors, which

© 2007 NRC Canada

Fig. 10. Longitude (*a*, *b*, *c*, *d*, and *e*) and latitude (*f*, *g*, *h*, *i*, and *j*) of five simulated tracks with varying observation window: 72 min towards the day and 14.4 min towards the night (*a* and *f*), 36 min towards the day and 14.4 min towards the night (*b* and *g*), 72 min towards the day and 7.2 min towards the night (*c* and *h*), 36 min towards the day and 7.2 min towards the night (*d* and *i*), and 18 min towards the day and 3.6 min towards the night (*e* and *j*). Each frame shows the true simulated track (solid line), the most probable track from the light-based state–space model (dots), and the 95% point-wise confidence region (shaded area). Finally, the estimated standard deviation (SD) of the most probable latitude is shown (*g*).


26 Dec. 2002 14 July 2003
Longitude
Latitude
SD(Latitude)
72 min. day, 14.4 min. night (a & f)
36 min. day, 14.4 min. night (b & g)
72 min. day, 7.2 min. night (c & h)
36 min. day, 7.2 min. night (d & i)
18 min. day, 3.6 min. night (e & j)
26 Dec. 2002
14 Feb. 2003
5 Apr. 2003
25 May 2003
14 July 2003
2 Sept. 2003
22 Oct. 2003
11 Dec. 2003
Date

are not all included in the model, the most obvious one being local weather conditions. If there is a heavy cloud cover one morning, but clear blue sky the next, the measurements from the latter will all be higher even though all other conditions are equal. The local weather conditions, or other factors influencing the light measurements, can change gradually, so the correlation structure further allows the correlation between two light measurements near the same solar event to depend on the elapsed time between them.

Another unique feature of the model presented here is that geolocations are estimated twice a day. Threshold algorithms need to know the solar altitude angle at two times to produce one geolocation, and obtain this from two solar events. Getting twice as many geolocations is naturally preferable, but the

© 2007 NRC Canada

underlying difference in model assumptions is even more important. Threshold algorithms assume that the movement of the tagged animal between two solar events can be neglected. The model presented here only assumes that movement in the approximately 90 min observational window around each solar event can be neglected. Furthermore, if there are substantial movements within the observational windows, the model will interpret that as additional serial correlation, which will correctly propagate into wider confidence regions.

The ability to estimate the yearly variation in the latitudinal standard deviation is an additional benefit of, and unique to, the model presented here. In Sibert et al. (2003) and Nielsen et al. (2006), an ad hoc parametric model had to be introduced to compensate for this pattern in threshold-based geolocations. It was difficult to verify that the assumed parametric structure was a valid approximation. The model presented here includes no assumptions about the yearly variation in latitudinal standard deviation, but the pattern clearly emerges as a simple consequence of propagating the data uncertainties through the geolocation process.

The two real data cases presented here show that the model presented is able to produce geolocations even in cases where the threshold-based latitudinal estimates are extremely inaccurate. The estimated most probable tracks are not perfect, but they are far superior to the threshold-based geolocations.

All model parameters can be identified from data. The estimated most probable track is able to follow the true track when realistic assumptions about movement pattern and observation uncertainty are made.

The model is surprisingly insensitive to varying the sampling rate of the light measurements. This implies that it is actually possible to get reliable geolocations from a small number (say 10) of correctly selected points around each solar event. This result is potentially very useful, as tags that transmit their data to satellites can only transfer a limited amount of data.

The fact that the model parameters can be identified correctly from data, and that the most probable tracks follow the true tracks, implies that approximation used in the unscented Kalman filter is sufficiently accurate for the nonlinearities in this model. This is fortunate, as the more accurate alternatives (simulation-based filters, e.g., Doucet et al. 2001) are far more computationally demanding.

Ideally, the light that is reflected by the moon should be part of the model, as it clearly influences the light measurements in the periods surrounding each full moon. The astronomical algorithms describing moon altitude and fraction of moon disc illuminated are straightforward to implement (Meeus 1998). Adding the two light sources (sun and moon) is the problematic part. Light measurements from the tag are processed by the inner workings of the tag. This processing can be viewed as a nonlinear transformation of the raw light intensities, and the details of this transformation are not easily obtainable, as the inner workings of some tags are not well documented or are simply considered proprietary. The asymmetric observation window chosen here avoids most of the complications with the moon, as it focus on the part of the light record where moonlight is negligible, but including the moon could potentially further improve this model.

The problems associated with standardizing light measurements taken at different depths, also referred to as depth-correcting, has been avoided in this paper, because the analyzed data were obtained at constant depths. The light measurements can be depth-corrected by the tag manufacturers before they are used in the model, or the model can be extended by a parametric model for the relationship between depth and light. It is expected that depth-correcting will introduce more uncertainty in the light measurements and consequently in the estimated tracks.

The suggestions to use narrow-band blue light instead (Ekstrom 2004) or to use cosine-collectors on the tags (Qayum et al. 2006) are interesting. If these light measurements are less sensitive to weather conditions or better capture light at depths, it will reduce the uncertainties and further improve the most probable track.

It has been shown that using sea surface temperature can improve geolocations (Teo et al. 2004; Nielsen et al. 2006). The model presented here can be extended in exactly the same way. However, instead of using sea surface temperature as a second patch (the first one being the Kalman filter) to a light-based algorithm that is giving extremely inaccurate geolocations, here it would only be used to further improve already reliable light-based geolocations.

Acknowledgements

The work described in this paper was sponsored by the University of Hawai'i Pelagic Fisheries Research Program under cooperative agreement No. NA17RJ1230 from the National Oceanic and Atmospheric Administration. We thank Kevin Wong (NOAA – Pacific Islands Fisheries Science Center) and Phil White, crew, and officers of the NOAA research ship *Townsend Cromwell* for deploying the drifter buoy and pop-up satellite archival tag, and Mike Musyl (NOAA – National Marine Fisheries Service) for providing the data. We thank Bruno Leroy from the Oceanic Fisheries Programme of the Secretariat of the Pacific Community for providing data from the New Caledonia mooring.

References

Arnold, G., and Dewar, H. 2001. Electronic tags in marine fisheries research: a 30-year perspective. *In* Electronic tagging and tracking in marine fisheries reviews: methods and technologies in fish biology and fisheries. *Edited by* J.R. Sibert and J.L. Nielsen. Kluwer Academic Press, Dordrecht. pp. 7–64.

Block, B.A. 2005. Physiological ecology in the 21st century: advancements in biologging science. Integr. Comp. Biol. **45**: 305–320.

Doucet, A., de Freitas, N., and Gordon, N. 2001. Sequential Monte Carlo methods in practice. Springer-Verlag, New York.

Ekstrom, P.A. 2004. An advance in geolocation by light. Mem. Natl Inst. Polar Res. Spec. Issue, **58**: 210–226.

Harvey, A.C. 1990. Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge.

Julier, S., Uhlmann, J., and Durrant-Whyte, H.F. 2000. A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Trans. Autom. Control, **45**(3): 477–482.

Meeus, J. 1998. Astronomical algorithms. Willmannn–Bell. Richmond, Va.

Musyl, M.K., Brill, R.W., Boggs, C.H., Curran, D.S., Kazama, T.K., and Seki, M.P. 2003. Vertical movements of bigeye tuna (*Thunnus obesus*) associated with islands, buoys, and seamounts near the main

© 2007 NRC Canada

Hawaiian Islands from archival tagging data. Fish. Oceanogr. **12**: 152–169.

Musyl, M.K., Brill, R.W., Curran, D.S., Gunn, J.S., Hartog, J.R., Hill, R.D., Welch, D.W., Eveson, J. P., Boggs, C. H., and Brainard, R.E. 2001. Ability of archival tags to provide estimates of geographical position based on light intensity. *In* Electronic tagging and tracking in marine fisheries reviews: methods and technologies in fish biology and fisheries. *Edited by* J.R. Sibert and J.L. Nielsen. Kluwer Academic Press, Dordrecht. pp. 343–368.

Nielsen, A., Bigelow, K.A., Musyl, M.K., and Sibert, J.R. 2006. Improving light-based geolocation by including sea surface temperature. Fish. Oceanogr. **15**: 314–325.

Qayum, H.A., Klimley, A.P., Newman, R., and Richert, J.E. 2006. Broad-band versus narrow-band irradiance for estimating latitude by archival tags. Mar. Biol. **151**: 467–481. doi:10.1007/s00227-006-0514-y.

Sibert, J., Musyl, M.K., and Brill, R.W. 2003. Horizontal movements of bigeye tuna (*Thunnus obesus*) near Hawaii determined by Kalman filter analysis of archival tagging data. Fish. Oceanogr. **12**: 141–151.

Smith, P., and Goodman, D. 1986. Determining fish movement from an "archival" tag: precision of geographical positions made from a time series of swimming temperature and depth. NOAA Tech. Memo. NOAA-TM-NMFS-SWFC-60.

Teo, S.L.H., Boustany, A., Blackwell, S., Walli, A., Weng, K.C., Block, B.A. 2004. Validation of geolocation estimates based on light level and sea surface temperature from electronic tags. Mar. Ecol. Prog. Ser. **283**: 81–98.

Wan, E.A., and van der Merwe, R. 2001. The unscented Kalman filter. *In* Kalman filtering and neural networks. *Edited by* S. Haykin. Wiley Publishing, New York. pp. 221–280.

Welch, D.W., and Eveson, J.P. 1999. An assessment of light-based geoposition estimates from archival tags. Can. J. Fish. Aquat. Sci. **56**: 1317–1327.

Wilson, S.G., Lutcavage, M.E., Brill, R.W., Genovese, M.P., Cooper, A.B., and Everly, A.W. 2005. Movements of bluefin tuna (*Thunnus thynnus*) in the northwestern Atlantic Ocean recorded by pop-up satellite archival tags. Mar. Biol. **146**: 409–423.

Wilson, S.G., Polovina, J.J., Stewart, B.S., and Meekan, M.G. 2006. Movements of whale sharks (*Rhincodon typus*) tagged at Ningaloo Reef, Western Australia. Mar. Biol. **148**: 1157–1166.

© 2007 NRC Canada

Context-dependent responses of juvenile Atlantic salmon (*Salmo salar*) to forestry activities at multiple spatial scales within a river basin

Julie Deschênes, Marco A. Rodríguez, and Pierre Bérubé

Abstract: We used classification trees and regression trees to relate the incidence and density of juvenile Atlantic salmon (*Salmo salar*) to forestry activities measured at four spatial scales (subbasin and 8, 2, and 0.5 km radii upstream of study sites) and environmental features in 120 stream reaches of the Cascapedia River basin, Québec, Canada. At all scales, incidence increased with reach size and accessibility to the reach from the river mainstem. Incidence declined with areal coverage of logging at all scales, but only in larger reaches. The time horizon over which logging effects were detected increased with spatial scale. At all scales, density in salmon-bearing reaches increased with accessibility. Density in more accessible reaches was negatively related to logging over the preceding 9 years at the subbasin and 8 km scales, but no effects of logging on density were detected at the 2 and 0.5 km scales. Overall, apparent effects of logging activities on salmon incidence and density were mostly negative and strong, but were both markedly scale-dependent and conditional on environmental context.

Résumé : Nous avons utilisé les arbres de classification et de régression pour relier l'incidence et la densité de saumons de l'Atlantique (*Salmo salar*) à l'environnement et aux activités forestières quantifiées à quatre échelles spatiales (le sous-bassin et des rayons de 8, 2 et 0,5 km en amont de chaque tronçon) dans 120 tronçons du bassin versant de la rivière Cascapédia, Québec, Canada. A toutes les échelles, l'incidence du saumon augmentait avec la taille des tronçons et l'accessibilité depuis la rivière principale et diminuait avec une augmentation du pourcentage de superficie coupée. Les effets des coupes étaient détectables sur des horizons temporels croissants avec l'augmentation de l'échelle spatiale. La densité de saumons augmentait avec l'accessibilité à toutes les échelles. Dans les tronçons plus accessibles, la densité diminuait avec une augmentation du pourcentage de superficie coupé au cours des 9 années précédentes aux deux plus grandes échelles du sous-bassin et de 8 km. Aucun effet des coupes sur la densité de saumons n'était détectable aux deux plus petites échelles de 2 et 0,5 km. Les effets apparents des coupes sur l'incidence et la densité de saumons juvéniles étaient forts et négatifs, mais dépendant de l'échelle spatiale et du contexte environnemental.

Introduction

The responses of aquatic organisms to human activities such as logging can be conditional on environmental context. For example, removal of riparian vegetation by logging in Oregon streams led to greater increase of periphyton and invertebrate abundance in small reaches than in large ones, apparently because riparian forest had greater influence on stream production in the small reaches (Murphy and Hall 1981). Logging also had a context-dependent effect on winter abundance of salmonid fry in Alaskan streams: fry abundance at logged sites declined if woody debris was removed, but increased otherwise (Murphy et al. 1986). Accounting for environmental context can lead to clearer understanding of the impacts of logging on stream fish populations and help develop preventive or remedial measures tailored to specific environmental conditions. Nonetheless, few studies have explicitly considered potential interactions between logging and the environmental context when assessing impacts of logging on fish populations (e.g., Murphy and Hall 1981; Murphy et al. 1986).

Logging activities affect stream fish populations by modifying key environmental features that structure these populations at various spatial scales. At the local habitat scale, removal of riparian canopy by logging can increase water temperature and light penetration (Davies and Nelson 1994), resulting in increased invertebrate abundance (Kiffney et al. 2003). At the drainage basin scale, logging can increase the load of suspended sediments and the magnitude of freshet discharge (Macdonald et al. 2003), as well as water yield

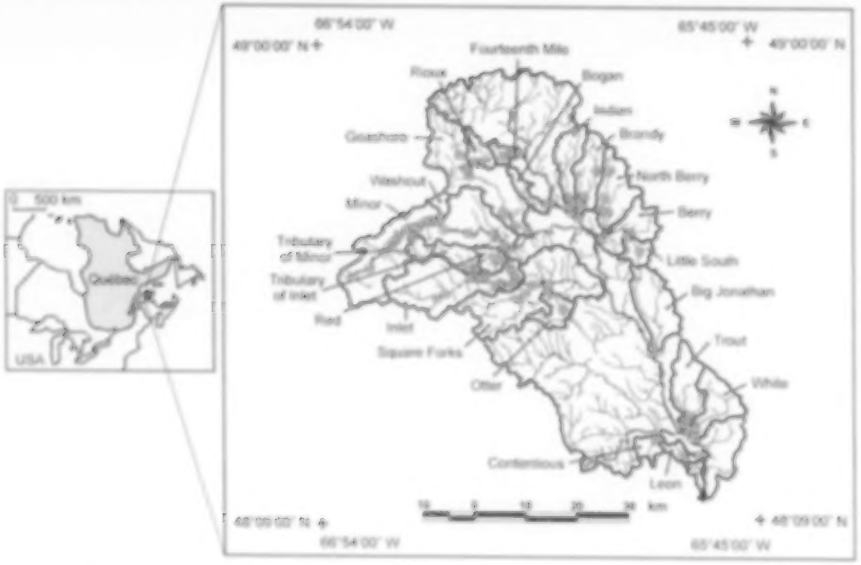
Received 6 July 2006. Accepted 12 April 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 3 August 2007. J19403

J. Deschênes and M.A. Rodríguez.[1] Département de chimie–biologie, Université du Québec à Trois-Rivières, C.P 500, Trois-Rivières, QC G9A 5H7, Canada.
P. Bérubé. Ministère des Ressources Naturelles et de la Faune, Direction de la Recherche sur la Faune, 930 Chemin Sainte-Foy, 4e étage, Québec, QC G1S 2L4, Canada.

[1]Corresponding author (e-mail: marco.rodriguez@uqtr.ca).

 doi:10.1139/F07-077 © 2007 NRC Canada

Fig. 1. Location of the 120 sampling sites distributed among 22 tributary streams in the Cascapedia River basin, Québec, Canada.



and nutrient concentrations (Lamontagne et al. 2000; Martin et al. 2000). Because fish relate to their local or proximal environment differently than to larger-scale characteristics of the basin (Armstrong et al. 1998; Folt et al. 1998), effects of logging on fish abundance may also differ across scales.

It is not clear which scale, local habitat or drainage basin, is most useful for understanding potential context-dependent effects of logging on fish populations. Logging has been suggested to have maximum impact at a scale of tens to hundreds of hectares (Carignan and Steedman 2000). However, basin-wide studies at larger scales may allow for more effective assessment of the cumulative effects of logging by integrating a multitude of small disturbances over a large basin (Hicks et al. 1991). Direct comparison of results and conclusions from studies conducted at different spatial scales can be misleading because patterns found at smaller scales do not necessarily extrapolate to larger scales (Wiens 1989; Hicks et al. 1991). Multiscale studies can therefore help identify the most appropriate spatial scales for understanding the impacts of logging activities on fish stream populations.

Logging effects on stream ecosystems have been examined mostly by comparing biotic or physical features of streams between logged and unlogged sites (e.g., Rowe et al. 2002; Death et al. 2003; Latterell et al. 2003). However, the areal coverage of logged patches within different age classes may be a more precise and informative measure of impact than simply the presence or absence of logging. For example, population densities of smaller fish species in Oregon streams declined with increases in coverage of recent (1–3 years) logging, whereas larger species declined with increase in coverage of older (4–8 years) logging (Rutherford et al. 1992).

This study examines responses of juvenile Atlantic salmon (*Salmo salar*) to forestry activities occurring 0–19 years before the study, measured at four spatial scales (subbasin and 8, 2, and 0.5 km radii upstream of study sites). Two biological metrics, incidence and population density in salmon-bearing reaches, were used as measures of population responses. Specifically, we compared the usefulness of proximate and larger-scale measurements of forestry activity for understanding the impacts of forestry on Atlantic salmon and examined whether those impacts were mitigated or accentuated by specific combinations of environmental conditions (i.e., whether forestry effects were context-dependent).

Materials and methods

The incidence and population density of Atlantic salmon and a host of environmental variables were quantified at 120 sampling sites distributed among 22 tributary streams in the Cascapedia River basin (3179 km^2), Québec, Canada (Fig. 1). The basin topography is hilly: 49% of the basin area has terrain slopes ≤15°; 28% has slopes >30°, and 23% has intermediate slopes. Plant cover consists primarily of fir–spruce forest (conifers, 66%; mixed, 23%; scattered deciduous, 2%), with the remaining area accounted for by shrubs, rocks, and other nonforest cover (9%; includes 2% in agricultural and urban land). The basin is sparsely populated but is subject to intensive forestry activities. Of the 120 sites in the present study, 116 were in subbasins exposed to logging in the 19 years preceding the study. Substantial portions of the subbasins upstream from the study sites (subbasin area: median = 70.1 km^2, range = 8.1–337.9 km^2) were affected by intensive logging (percentage of upstream area covered by cuts: median = 18.0%, range = 0.0%–37.6%) over this period. Governmental regulations have required forestry operations in the basin to leave an unharvested 20 m buffer strip (60 m when Atlantic salmon is present) along permanent streams since 1988.

Sites were visited in random sequence during low flow from mid-June to late August in 2000 (24 sites), 2001 (48 sites), and 2002 (48 sites). Sampled areas covered the entire

© 2007 NRC Canada

Table 1. Median, interquartile range, and variable codes for 25 environmental variables describing fish habitat at 120 sites in the Cascapedia River basin.

Variable name	Code	Median	Quartiles, 25%–75%
Accessibility index	AI	0.8	0.3–1.0
Altitude (m)	AL	240.0	167.5–315.0
Overhead opening (°)	OO	95.7	75.4–113.5
Cover index	CI	2.4	1.4–3.0
Distance to mainstem (km)	DM	9.1	2.5–21.7
Distance to mouth (km)	DO	75.5	52.8–82.4
Entrenchment (%)	EN	10.0	5.0–15.0
Height increment at flood (m)	HI	0.5	0.4–0.7
Mean current velocity ($cm \cdot s^{-1}$)	MC	34.9	26.0–48.9
Mean depth (cm)	MD	24.6	18.8–29.4
Mean substratum size	MS	5.3	4.9–5.8
Mean temperature (°C)	MT	10.5	9.0–12.0
Mean wetted width (m)	MW	9.3	6.3–14.1
Pools (number)	PO	0.0	0.0–1.5
Stream gradient (°)	SG		
Subbasin		2.1	0.6–3.3
8 km		1.1	0.5–2.3
2 km		1.0	0.5–1.7
0.5 km		1.1	0.3–2.0
Stream order	SO	3.0	3.0–4.0
Stream slope (°)	SS	1.0	0.6–1.3
Subbasin area (km^2)	SA	70.1	25.5–151.0
Terrace width[a] (m)	TW	107.8	81.3–163.5
Units of large woody debris (number)	UW	3.5	1.0–8.5
Vegetation abundance index	VA	1.0	1.0–1.8
Width increment at flood[a] (m)	WI	2.6	1.9–4.0

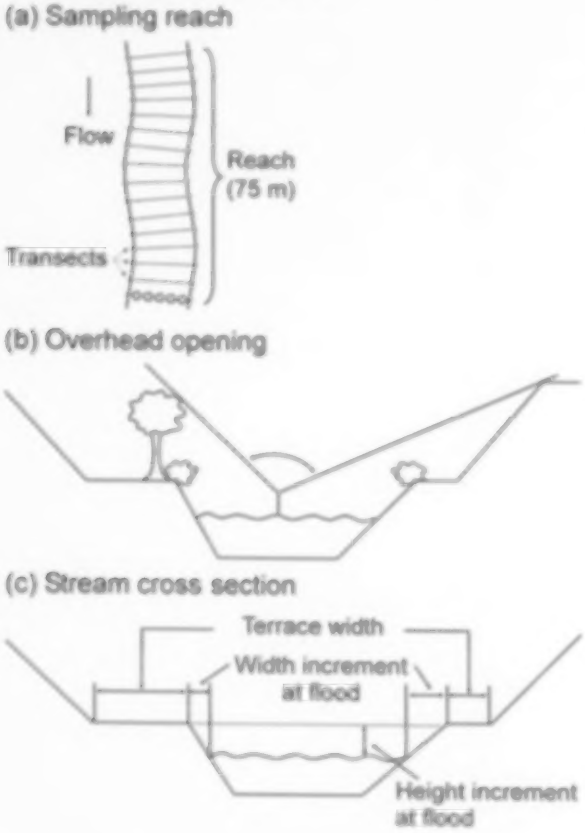
Note: Variable codes for Fig. 4 are also given.
[a]Sum of measures from right and left margins.

stream width in completely wadable reaches and ranged 5 m from one bank, chosen randomly, towards the opposite bank in deeper reaches. Fish were sampled by single-pass electrofishing (Smith-Root D-15) in an upstream direction within open stream sections (Lobón-Cerviá and Utrilla 1993; Jones and Stockwell 1995). All captured fish were identified, measured, weighed, and returned to the stream site of capture. Salmon density was calculated as total captures divided by section area (numbers·100 m^{-2}). Because capture efficiency was not 100%, this measure underestimates true density but should be proportional to it if efficiency is comparable across reaches. All reaches were sampled at base flow, and therefore variation in depth, water velocity, cover, and substratum size among reaches was limited (Table 1). Abundance estimates obtained by one-pass electrofishing have been shown to correlate well with those from multiple-pass removal experiments in homogeneous geographic areas (e.g., small mountain streams (Kruse et al. 1998); small warmwater streams (Edwards et al. 2003)), including estimates for juvenile Atlantic salmon in areas similar to those sampled in the present study (Hedger et al. 2005; M.A. Rodríguez, unpublished data).

In all, 25 environmental variables were quantified (Table 1). Water depth and substratum size were recorded at five equidistant points along each of 16 transects perpendicular to streamflow (Fig. 2*a*). Substratum size was assigned a numerical code reflecting the dominant size category (modified Wentworth scale: 1, ≤0.0625 mm; 2, 0.0625–2 mm; 3, 2–16 mm; 4, 16–32 mm; 5, 32–64 mm; 6, 64–256 mm; 7, >256 mm; 8, bedrock). Current velocity (pigmy-type meter, Scientific Instruments 1205) was measured at five equidistant points along every third transect starting with the second transect from the downstream end. Wetted width was measured at each transect. Abundance of submerged vegetation (moss or macrophytes) and overall availability of structural cover (rocks, woody debris, undercut bank, and overhanging vegetation) were estimated visually and assigned ordinal values reflecting areal coverage (1, ≤5%; 2, 6%–15%; 3, 16%–45%; 4, >45%). Overhead opening (angle between riparian canopy or hilltops blocking incident sunlight at the centre of the stream; Fig. 2*b*) and slope over the stream reach were measured with a handheld clinometer (Suunto PM-5). The increment in stream height and width at flood (from annual flood marks) and terrace width (distance between stream bank and piedmont) were measured on site (Fig. 2*c*). Water temperature (midpoint of maximum and minimum recorded while electrofishing) was measured (maximum–minimum thermometer, Barigo). Entrenchment (mean gradient ≤100 m away from stream bank), subbasin area, stream order (Strahler), and altitude were obtained from 1:20 000 topographic maps, as were distances by waterway from each reach to the Cascapedia River ("distance to mainstem") and from the mouth of each stream to the mouth of the Cascapedia River ("distance to mouth")

© 2007 NRC Canada

Fig. 2. Schematic views illustrating field measurements of (*a*) depth and substratum size, recorded at sampling points (circles) along all transects, and current velocity, recorded at sampling points along every third transect starting from the second-most downstream transect; (*b*) overhead opening (angle between structures blocking direct sunlight); (*c*) terrace width, height increment at flood, and width increment at flood.

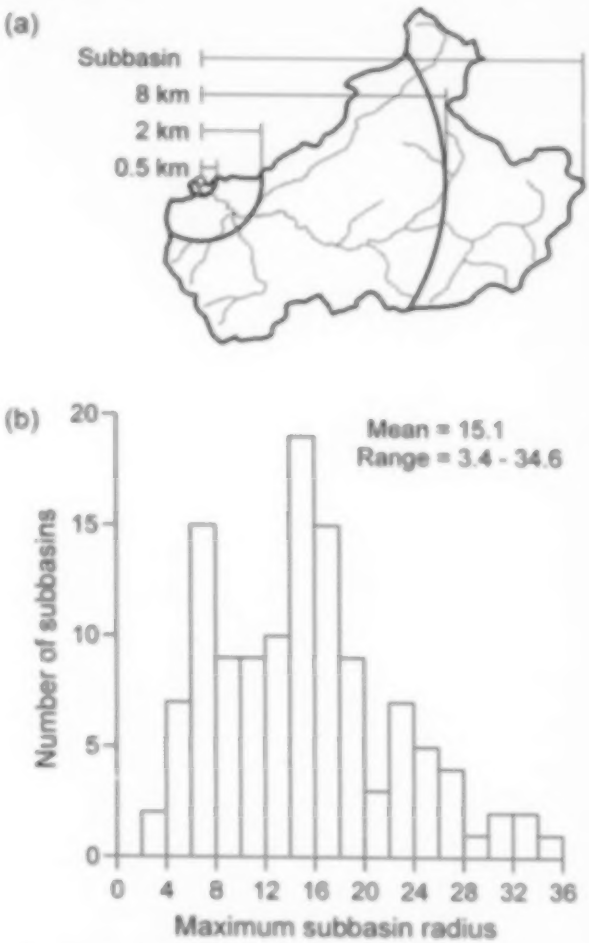


(planimeter, Calculated Industries 6125). Units of large woody debris (>10 cm diameter) and pools were counted within each reach. Physical barriers potentially affecting upstream migration of fish along a tributary were assessed from field observations and topographic maps, and their effectiveness was coded as an integer value, B, ranging from 0 (no visible barrier) to 4 (insurmountable barrier) and reflecting the height, type (beaver dams, log jams, culverts, falls), and configuration of the barrier. An index of accessibility combining multiplicatively all potential barriers for each site was calculated as

$$\text{accessibility} = \prod_{i=1}^{N}\left(1 - \frac{B_i}{4}\right)$$

where N is the number of visible barriers and B_i is the effectiveness of barrier i downstream from the site. Accessibility was assigned the value 1 in the absence of visible barriers. The index thus ranged from 0 to 1, taking a value of 0 if at least one barrier was insurmountable ($B_i = 4$).

Fig. 3. (*a*) Schematic representation of zones used to quantify landscape features and forestry activities at four spatial scales: subbasin and 8, 2, and 0.5 km radii upstream of the sampling sites. The concentric curves are delimited by watershed boundary lines. (*b*) Frequency distribution of the radius from the study reach to the most distant upstream point within the subbasin of the reach (km).



A geographic information system (ArcGIS, v. 8.2, ESRI, Redlands, California) was used to measure logging activities at four spatial scales: subbasin and semicircular zones of 8, 2, and 0.5 km radii upstream of each site (Fig. 3*a*), as well as the total subbasin area upstream of the site, and stream gradient (mean slope from the site to upstream boundary), the percentage of the subbasin area affected by spruce budworm outbreaks, and the density of roads (primary, secondary, tertiary, and total; km·km^{-2}). GIS maps were based on 1:20 000 forest maps (Ministère des Ressources naturelles du Québec), annual logging records, and photo-interpretation (Société de la faune et des parcs du Québec). All but 25 of the study reaches had subbasins that extended beyond the 8 km radius (Fig. 3*b*). Logging activities were grouped into three categories reflecting the harvesting procedures employed by forestry companies: intensive (clearcut harvesting, protection of

© 2007 NRC Canada

Table 2. Median and interquartile range for 35 forestry variables measured at four spatial scales in the Cascapedia River basin.

Variable name	Code	Spatial scale			
		Subbasin	8 km	2 km	0.5 km
Spruce budworm outbreaks	sb	22.6 (18.8–27.8)	25.8 (17.3–32.2)	21.3 (6.5–37.8)	9.7 (0.0–23.4)
Road density					
Total	rt	1.9 (1.4–2.3)	1.8 (1.5–2.1)	1.5 (1.0–2.2)	1.6 (0.6–2.8)
Primary	r1	0.2 (0.1–0.3)	0.1 (0.0–0.3)	0.0 (0.0–0.4)	0.0 (0.0–0.9)
Secondary	r2	0.8 (0.2–1.6)	0.4 (0.1–1.2)	0.2 (0.0–0.8)	0.0 (0.0–0.0)
Tertiary	r3	0.6 (0.2–1.0)	0.8 (0.3–1.4)	0.8 (0.2–1.4)	0.0 (0.0–1.6)
Intensive logging (years)					
0–1	i1	1.0 (0.0–2.8)	0.3 (0.0–4.6)	0.0 (0.0–2.3)	0.0 (0.0–0.0)
0–2	i2	2.5 (0.0–3.5)	1.9 (0.0–5.4)	0.0 (0.0–5.8)	0.0 (0.0–0.0)
0–3	i3	2.8 (0.7–4.3)	2.4 (0.0–5.6)	0.0 (0.0–9.0)	0.0 (0.0–2.2)
0–4	i4	3.1 (0.8–5.6)	2.3 (0.0–7.9)	0.0 (0.0–10.0)	0.0 (0.0–4.0)
0–9	i9	8.6 (2.9–14.2)	5.1 (2.2–18.9)	2.4 (0.0–24.5)	0.0 (0.0–5.4)
0–14	i14	13.1 (4.0–22.4)	11.4 (4.0–23.8)	6.3 (0.3–25.1)	0.0 (0.0–11.5)
0–19	i19	16.4 (4.5–23.8)	17.8 (4.8–25.6)	8.7 (1.3–25.5)	0.0 (0.0–17.4)
Partial logging (years)					
0–1	p1	0.1 (0.0–2.2)	0.0 (0.0–0.3)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
0–2	p2	0.6 (0.0–3.1)	0.1 (0.0–1.6)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
0–3	p3	0.9 (0.0–3.9)	0.1 (0.0–2.8)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
0–4	p4	1.5 (0.0–5.8)	1.0 (0.0–5.1)	0.0 (0.0–0.5)	0.0 (0.0–0.0)
0–9	p9	3.8 (0.8–8.5)	2.3 (0.1–6.9)	0.0 (0.0–3.1)	0.0 (0.0–0.0)
0–14	p14	6.1 (0.9–11.1)	4.3 (0.1–9.0)	0.0 (0.0–6.7)	0.0 (0.0–0.0)
0–19	p19	7.2 (0.9–13.4)	4.9 (0.1–9.1)	0.0 (0.0–7.2)	0.0 (0.0–0.3)
Combined logging (years)					
0–1	c1	2.7 (0.3–5.5)	2.7 (0.0–6.3)	0.0 (0.0–7.1)	0.0 (0.0–1.5)
0–2	c2	3.9 (2.6–7.2)	4.0 (2.1–8.4)	0.0 (0.0–13.2)	0.0 (0.0–3.6)
0–3	c3	5.1 (2.9–8.9)	5.5 (2.5–9.4)	0.0 (0.0–16.2)	0.0 (0.0–4.4)
0–4	c4	7.8 (3.7–12.5)	5.8 (3.3–13.4)	5.1 (0.0–17.1)	0.0 (0.0–9.7)
0–9	c9	15.4 (6.7–23.0)	10.6 (5.7–22.6)	8.8 (0.6–31.1)	0.0 (0.0–11.7)
0–14	c14	18.3 (10.3–29.5)	18.5 (7.3–30.0)	12.6 (1.8–31.9)	3.9 (0.0–21.4)
0–19	c19	21.5 (11.1–35.4)	23.6 (7.7–31.2)	18.7 (2.9–34.5)	4.4 (0.0–24.5)

Note: Road densities are expressed as km·km^{-2}. Spruce budworm outbreaks and logging data are given as percent of total subbasin area. Variable codes for Fig. 4 are also given.

regeneration, seed tree, and patch cutting), partial (partial harvesting, thinning, strip harvesting, selection cutting, harvesting within green strips, and diameter-limit cutting), and combined (intensive plus partial) logging. The logging variables used in the analyses quantified the percentage of the total land area at a given distance upstream that was logged (either intensively, partially, or combined) in the following time intervals preceding the study: 0–1, 0–2, 0–3, 0–4, 0–9, 0–14, and 0–19 years (Table 2).

We used classification trees (CT) and regression trees (RT) to examine the context-dependent response of salmon populations to forestry variables measured at multiple scales. CT and RT are robust tools for ecological modelling that can effectively represent nonlinear relationships between a single response variable (qualitative in CT and quantitative in RT) and a set of predictor variables. Additionally, they can reveal nonadditive effects and complex interactions between predictors without requiring a priori specification of the form of these interactions (Breiman et al. 1984; De'ath and Fabricius 2000). This ability to detect interactions is particularly relevant in studies focusing on context-dependent effects, which by definition are interactive. CT and RT use binary recursive partitioning (Breiman et al. 1984) to split the data into progressively smaller and more homogeneous subgroups. Beginning with the entire data set, at each step this algorithm examines all possible splits for each possible value of the predictor variables and selects the candidate split (the "splitting value") that maximizes the homogeneity within the two resulting subgroups with respect to the response variable. Similar to parametric methods such as multiple regression, in CT and RT strong colinearity among predictor variables can complicate the interpretation of effects of individual predictors. However, in CT and RT only one predictor is allowed to enter the model at any given split, and as the data are split into smaller subgroups, the relations among predictors may change. Therefore, predictors that are highly correlated over the whole data set may not be as strongly correlated in subgroups of the data (Lawler and Edwards 2002).

Salmon incidence and density were treated separately because salmon were absent from 51 of the 120 reaches. CT were used to analyse salmon incidence in all 120 reaches, whereas RT were used to analyse salmon density only in the 69 salmon-bearing reaches. Although RT are invariant to the

© 2007 NRC Canada

distribution of predictor variables, they are not invariant to the distribution of the response variable (Atkinson and Therneau 2000). Salmon density was therefore transformed as $\ln(X)$ to reduce the potential influence of outliers. CT and RT were produced for each spatial scale separately.

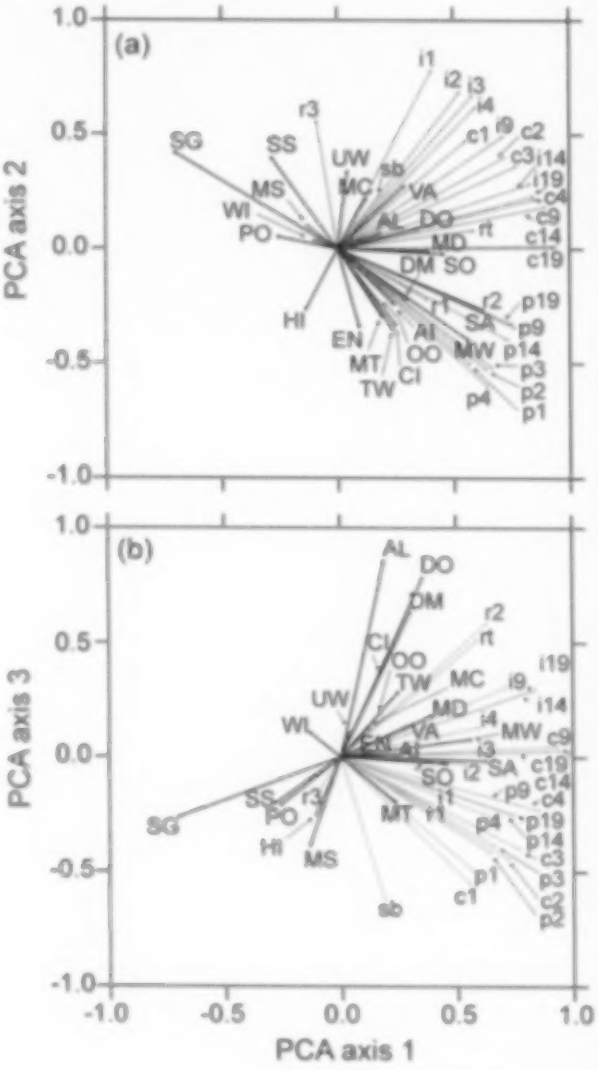
Model specification and cross-validation followed Turgeon and Rodríguez (2005). The Rpart software library (Atkinson and Therneau 2000) was used for all calculations (S-PLUS program, v.7.2, Insightful Corporation, Seattle, Washington). Because recursive partitioning sometimes leads to excessive splitting, producing trees that are too large and overfit the data, we used "pruning" and cross-validation (an alternative to conventional parametric inference for model assessment) to select optimal trees (Atkinson and Therneau 2000; De'ath and Fabricius 2000). First, we generated a sequence of trees of increasing size, using a cost-complexity parameter (CP), to eliminate ("prune off") splits that were obviously not worthwhile, i.e., that did not improve the fit by at least the value of CP (= 0.01 for all trials) (Atkinson and Therneau 2000). Then, 10-fold cross-validation was used to estimate prediction error, and final tree size was determined by the 1 SE rule, which favours the largest tree for which the cross-validated error falls within 1 standard error (SE) of the minimum relative error determined by cross-validation (Atkinson and Therneau 2000). The influence of individual predictor variables was gauged by the proportional reduction in error (PRE; a measure of the variation accounted for by the splits associated with each predictor in the tree), an approach similar to the use of partial R^2 to assess the contribution of individual predictors in multiple regression. Additionally, we used the correct classification rate (CCR; percentage of all cases correctly predicted), sensitivity (percentage of observed presences correctly predicted), and specificity (percentage of observed absences correctly predicted) to evaluate the accuracy of CT. Empirical odds ratios, calculated as

$$\mathrm{OR} = \frac{P_1 A_2}{P_2 A_1}$$

where P_x and A_x represent numbers of reaches where salmon was present (P) or absent (A) under condition x, were used to compare salmon incidence under different environmental conditions determined by the CT, e.g., in reaches with wetted width <8.7 m vs. reaches with wetted width ≥8.7 m. A small constant (0.5) was added to values of P and A to avoid null denominators.

We conducted a preliminary principal components analysis (PCA; based on the correlation matrix) to examine the relationships between the environmental and forestry variables (measured at the subbasin scale) and check for multicolinearity among these variables. Variables were transformed as required to reduce the influence of outliers and linearize relationships. We did not retain a reduced set of components from the PCA for subsequent inclusion as predictors in CT and RT because of the difficulties in interpreting composite variables and because using linear composites of variables as predictors can reduce predictive power and mask interactions among the original variables (Muñoz and Felicísimo 2004).

Fig. 4. Loadings from a principal component analysis (PCA) of environmental (solid arrows, uppercase codes) and forestry (shaded arrows, lowercase codes) variables measured at the subbasin scale. Plots are presented for (*a*) axes 1 (30.6% of total variance) and 2 (13.8%) and (*b*) axes 1 and 3 (11.6%). Linear correlations between variables are represented approximately by the length and alignment of their respective arrows. Variable codes are as in Tables 1 and 2.



Results

The PCA of environmental and forestry variables showed little relationship between the areal coverages of intensive and partial logging in the Cascapedia River basin (Fig. 4). Combined logging, given by the sum of intensive and partial logging, was positively related to both these variables, as would be expected. Within each of these three sets of logging variables, correlations reflected the overlap in temporal coverage between pairs of variables, e.g., strong correlation

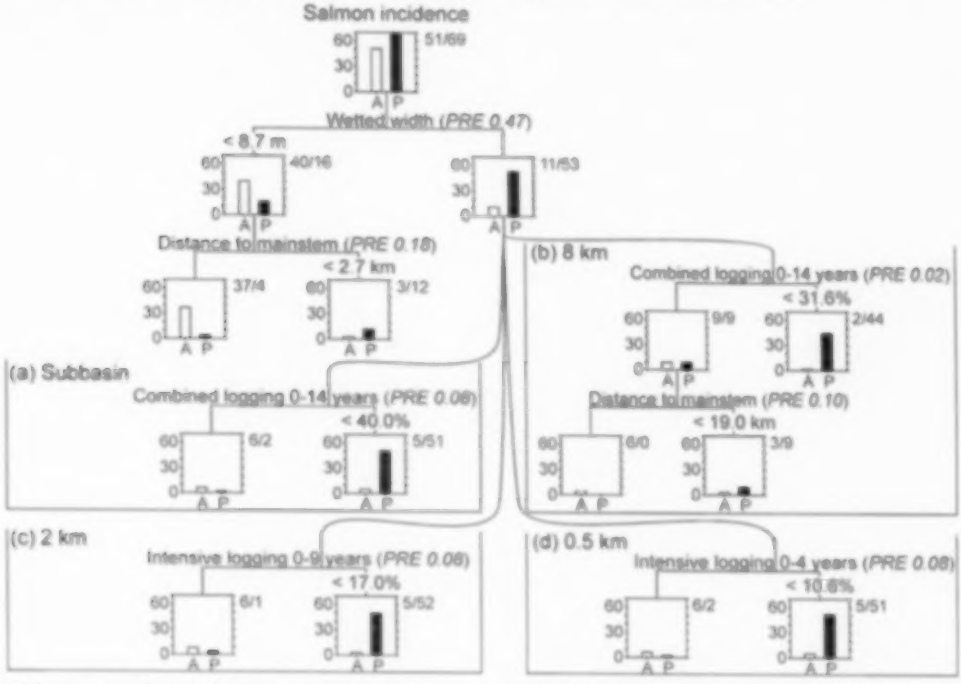
© 2007 NRC Canada

Table 3. Classification success (CCR, specificity, and sensitivity) and proportional reduction in error (PRE) of classification tree models and PRE of regression tree models by spatial scale.

	Classification trees				Regression trees
Spatial scale	CCR	Specificity	Sensitivity	PRE	PRE
Subbasin	88.3	84.3	91.3	72.5	68.7
8 km	90	84.3	94.2	76.5	68.4
2 km	89.2	84.3	92.8	72.5	57.5
0.5 km	88.3	84.3	91.3	72.5	49.1

Note: All values are percentages.

Fig. 5. Classification tree models for salmon incidence. Models are composed of a common part (top, left) and a part specific to each scale (open boxes; bottom): (*a*) subbasin, (*b*) 8 km, (*c*) 2 km, and (*d*) 0.5 km. Vertical bars represent the number of reaches with (solid) or without (open) salmon at each node. Splitting rules and proportional reduction in error (PRE) are given on the branches of each tree. The number of reaches with and without salmon are given alongside the boxes for each node.



between 0- to 14-year and 0- to 19-year cuts (14-year overlap), but weaker correlation between 0- to 14-year and 0- to 1-year cuts (1-year overlap). Among the environmental variables, the strongest correlations were between wetted width and subbasin area (Pearson $r_p = 0.86$), altitude and distance to mouth ($r_p = 0.84$), altitude and distance to mainstem ($r_p = 0.77$), stream gradient and subbasin area ($r_p = -0.74$), and stream gradient and wetted width ($r_p = -0.70$) (Fig. 4); all other correlations were ≤0.63. The strongest correlation between environmental and logging variables was between stream gradient and 0- to 19-year combined cuts ($r_p = -0.62$). Logging variables were generally related positively to subbasin size and wetted width and negatively to stream gradient (Fig. 4), perhaps reflecting greater accessibility and ease of manoeuvring for forestry equipment in large expanses of flat lowland terrain.

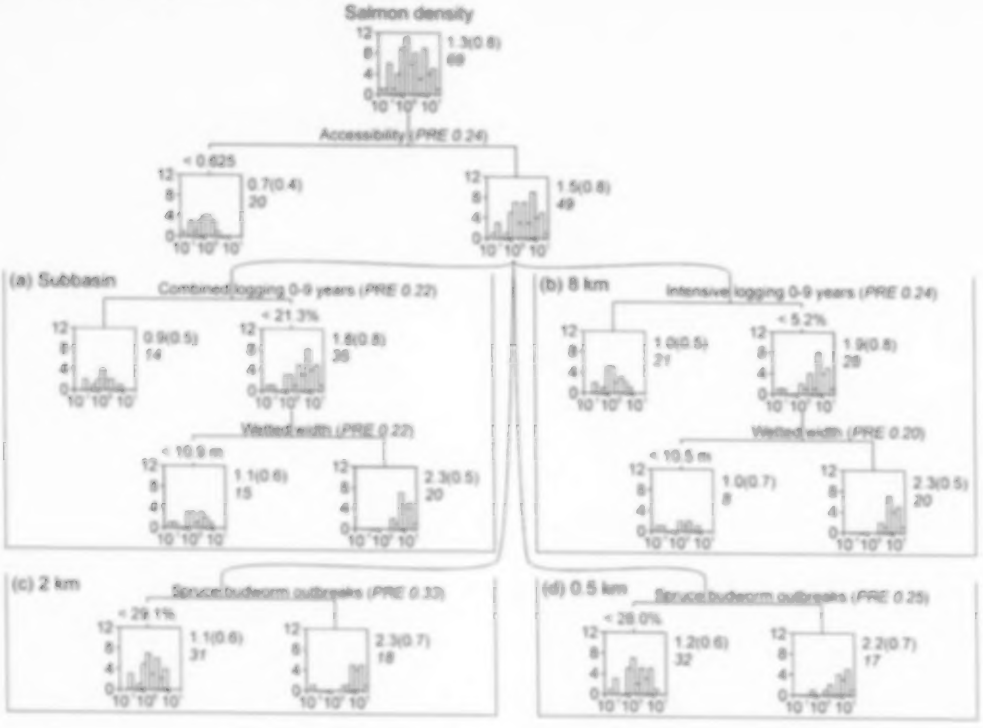
A total of 1539 salmon were captured over the three years of sampling. Salmon were present in 69 of the 120 reaches and 14 of the 22 streams sampled. The ability of final CT models to classify reaches correctly on the basis of incidence (CCR) differed little across scales and ranged from 88.3% at the subbasin scale to 90.0% at the 8 km scale (Table 3). All models had a greater ability to correctly classify salmon-bearing reaches (sensitivity) than reaches without salmon (specificity). For CT models, the proportional reduction in error varied little across scales (72.5%–76.5%).

Although CT models were built for each spatial scale separately, two environmental predictors, wetted width and distance to the mainstem, were retained in the final models at all scales (Fig. 5). Smaller reaches (wetted width <8.7 m) had lower incidence than larger reaches (odds ratio (OR) = 11.4). Among smaller reaches, salmon incidence was greater in reaches nearer (<2.7 km) to the Cascapedia River mainstem than in more distant reaches (OR = 29.8) (Fig. 5).

Among larger reaches (wetted width ≥8.7 m), salmon incidence was negatively related to logging at all scales

© 2007 NRC Canada

Fig. 6. Regression tree models for population density (numbers·100 m^{-2}) in salmon-bearing reaches (*X* axis). Models are composed of a common part (top) and a part specific to each scale (open boxes, bottom): (*a*) subbasin, (*b*) 8 km, (*c*) 2 km, and (*d*) 0.5 km. Vertical bars represent the density distribution of salmon at each node. Splitting rules and proportional reduction in error (PRE) values are given on the branches of each tree. The mean (SD) of ln-transformed density and the number of observations (in italics) are given alongside the boxes representing terminal nodes.



(Figs. 5*a*–5*d*). Salmon incidence was negatively related to combined logging in the preceding 14 years at both the subbasin (OR = 24.3) and 8 km scales (OR = 17.8) (Figs. 5*a*, 5*b*). Salmon incidence also was negatively related to intensive logging in the preceding 9 years at the 2 km scale (OR = 41.4) and to intensive logging in the preceding 4 years at the 0.5 km scale (OR = 24.3) (Figs. 5*c*, 5*d*). At the two smaller scales, declines in salmon incidence were apparent at relatively low threshold values for intensive logging (17.0% at the 2 km scale; 10.6% at the 0.5 km scale). The fact that salmon incidence was negatively related to logging only in larger reaches points to an interaction between logging and stream size that is independent of scale (Fig. 5).

Final RT models for salmon density in salmon-bearing reaches retained only two or three predictor variables. The proportion of variation in salmon density accounted for by the models increased with spatial scale, ranging from 49.1% at the 0.5 km scale to 68.7% at the subbasin scale (Table 3). Accessibility was a major predictor of salmon density at all scales (Fig. 6). Mean salmon density in readily accessible reaches (accessibility ≥0.625) was approximately twice that in less accessible ones (Fig. 6).

Relationships between salmon density and logging were only detected in readily accessible reaches, at the two largest spatial scales (Fig. 6). At the subbasin scale, mean salmon density in reaches having <21.3% of combined logging in the preceding 9 years was 2.5 times that in reaches with greater values of combined logging (Fig. 6*a*). Among the former reaches, those having wetted width ≥10.9 m had greater salmon density than narrower ones. At the 8 km scale, mean salmon density in reaches having <5.2% of intensive logging in the preceding 9 years was 2.5 times that in reaches with greater values of intensive logging (Fig. 6*b*). At this scale, mean salmon density was greatest in reaches having <5.2% of intensive logging in the preceding 9 years and wetted width ≥10.5 m. Final RT models were similar for the 2 km and 0.5 km scales. At these scales, salmon density was greatest in readily accessible reaches subject to spruce budworm outbreaks having areal coverage >28%–29.1% (Figs. 6*c*, 6*d*). Similar to the models for incidence, an interaction involving logging and an environmental predictor was apparent. Salmon density was negatively related to logging only in readily accessible reaches but, contrary to the models for incidence, only at the two largest spatial scales (Figs. 6*c*, 6*d*).

Discussion

At all spatial scales, logging appeared to have context-dependent effects: incidence declined only in large reaches and density declined only in readily accessible reaches. Logging impacts on periphyton, invertebrates, and physical characteristics in Oregon streams varied as a function of stream width and gradient and resulted in greater abundance

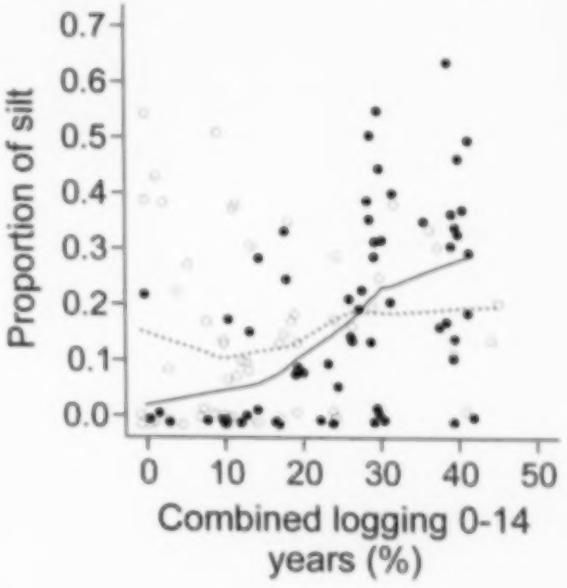
© 2007 NRC Canada

of periphyton and invertebrates in small, high-gradient reaches than in large, low-gradient ones (Murphy and Hall 1981). Canopy removal resulted in a smaller reduction in shading, and thus less enhancement of primary production, in large reaches than in small reaches. Furthermore, sediment deposition was greater in large reaches because of their lower gradient and transport capacity relative to small reaches. Increased sedimentation offset any beneficial effect of increased light exposure in larger reaches (Murphy and Hall 1981). In the Cascapedia River basin, the effect of logging on salmon incidence in large reaches may have been linked to decline in habitat quality mediated by increased deposition of fine sediments. Because reach width and gradient were negatively correlated (Spearman rank correlation, $r_s = -0.66$), small reaches likely were less susceptible to accumulation of fine sediments than large ones, possibly explaining why no effect of logging on salmon incidence was apparent in small reaches. The proportion of silt increased with logging in larger reaches but not in smaller ones (analysis of covariance (ANCOVA) term for the interaction between stream width and logging, $p = 0.007$; Fig. 7), lending support to this interpretation. The apparent context-dependent effect of logging on salmon density may be a consequence of restriction of fish movement by barriers such as log jams, beaver dams, culverts, or waterfalls in less accessible reaches, which may hinder adjustment of local densities to match spatial differences in habitat quality induced by logging. In contrast, salmon in accessible reaches may move more freely and avoid areas impacted by logging, creating density differentials that better reflect habitat quality.

The cumulative effects of logging on salmon incidence seemed to be integrated over longer time horizons at larger spatial scales. Fine sediments produced by logging have been reported to rapidly increase in reaches immediately downstream of the logged area and subsequently decline within 1 to 3 years after the disturbance (Macdonald et al. 2003; Haggerty et al. 2004). However, redistribution and transport of deposited sediments to reaches further downstream may continue for many years. For example, in the Big Beef Creek basin (USA), mass influx of sediments arising from forestry activities between 1850 and 1930 produced a sediment wave that was transported over 18 km of main stream channel and an artificial lake (≈1 km in length) (Quinn and Peterson 1996) before arriving near the downstream limit of the creek by 1970 (Madej 1982). This wave was then transported over a distance of 6 km between 1970 and 1976 (Madej 1982). Logging effects on downstream reaches may therefore be detectable for many decades after the actual disturbance (e.g., Madej 1982; Swank et al. 2001). Transport of fine sediments from upstream to downstream reaches over a span of decades may explain why logging effects appeared to be integrated over longer time periods at larger spatial scales. These long-term effects seem akin to the "ghost of the land use past", where past agricultural activity in North Carolina basins determined present-day fish and invertebrate diversity, independently of riparian reforestation (Harding et al. 1998).

In contrast with salmon incidence, logging effects on salmon density were only apparent at the two largest spatial scales. When fine sediments from multiple headwater sources accumulate in downstream reaches, effects of logging on fish populations may be magnified in those reaches that receive the impact of distant upstream sources and thus be apparent only at larger spatial scales. Salmon density was positively related to areal coverage of spruce budworm outbreaks at the two smaller scales, 2 km and 0.5 km. Loss of riparian canopy caused by defoliation can result in increased light exposure, local temperature, and productivity (Swanston 1991; Kiffney et al. 2003, 2004). The influence of defoliation on salmon density seems to be local (Swanston 1991) and may thus be more readily detectable at smaller spatial scales.

Fig. 7. Relative abundance of silt (particle size <0.0625 mm; transformed as arcsine square-root) vs. percent areal coverage of 0–14 years of combined logging, at the subbasin scale. Circles represent reaches in two categories of wetted width derived from CT models (Fig. 4): ≥8.7 m (solid circles) and <8.7 m (open circles). Lines are lowess regression curves for the two categories: ≥8.7 m (solid line) and <8.7 m (dotted line). A small amount of random jitter was added to improve the visibility of overlapping symbols.



In summary, apparent effects of logging activities on salmon incidence and density were mostly negative and strong, but were both markedly scale-dependent and conditional on environmental context. The two biological metrics used to quantify the responses to logging, incidence and density, differed in their sensitivity to spatial scale, capacity to detect logging impacts, and integration of effects over time. Examining the variation in salmon incidence and density over multiple spatial scales helped us understand the context-dependent responses of salmon populations to logging and pointed to potential limitations of studies focusing on single scales and biological metrics. For example, the apparent absence of logging effects on salmon density at the two smaller spatial scales suggests that future study designs should consider the possibility that large-scale disturbances arising from logging may not be detectable at smaller spatial

© 2007 NRC Canada

or temporal scales. Potential effects of logging should therefore be examined with a view to detection of downstream effects, particularly as concerns transport of sediments from headwaters to downstream reaches. Because biotic responses to logging can differ across geomorphic settings, future studies may benefit from considering environmental context explicitly when assessing the effects of logging on fish populations.

Acknowledgements

We thank V. Bérard, K. Brassard, C. Cossette, M.-J. Gagnon, E. Racicot, and C. Sauvé for expert field assistance, M.-A. Bernard and M. Gauthier for logistic support, and M. Lapointe for advice on site selection. Financial support was provided by the Natural Sciences and Engineering Research Council of Canada, le Fonds Québécois de la recherche sur la nature et les technologies, la Société Cascapédia, le Centre interuniversitaire de recherche sur le saumon atlantique (CIRSA), and le Syndicat des professeurs et professeures de l'UQTR. This paper is a contribution to the CIRSA program.

References

Armstrong, J.D., Grant, J.W.A., Forsgren, H.I., Fausch, K.D., DeGraaf, R.M., Fleming, I.A., Prowse, T.D., and Schlosser, I.J. 1998. The application of science to the management of Atlantic salmon (*Salmo salar*): integration across scales. Can. J. Fish. Aquat. Sci. **55**: 303–311.

Atkinson, E.J., and Therneau, T.M. 2000. An introduction to recursive partitioning using the Rpart routines. Mayo Foundation, Rochester, New York. Technical Rep. No. 61.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. 1984. Classification and regression trees. Chapman and Hall, New York.

Carignan, R., and Steedman, R.J. 2000. Impacts of major watershed perturbations on aquatic ecosystems. Can. J. Fish. Aquat. Sci. **57**(S2): 1–4.

Davies, P.E., and Nelson, M. 1994. Relationships between riparian buffer widths and the effects of logging on stream habitat, invertebrate community composition and fish abundance. Aust. J. Mar. Freshw. Res. **45**: 1289–1305.

De'ath, G., and Fabricius, K.E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology, **81**: 3178–3192.

Death, R.G., Baillie, B., and Fransen, P. 2003. Effect of *Pinus radiata* logging on stream invertebrate communities in Hawke's Bay, New Zealand. N.Z. J. Mar. Freshw. Res. **37**: 507–520.

Edwards, M.R., Combs, D.L., Cook, S.B., and Allen, M. 2003. Comparison of single-pass electrofishing to depletion sampling for surveying fish assemblages in small warmwater streams. J. Freshw. Ecol. **18**: 625–634.

Folt, C.L., Nislow, K.H., and Power, M.E. 1998. Implications of temporal and spatial scale for Atlantic salmon (*Salmo salar*) research. Can. J. Fish. Aquat. Sci. **55**: 9–21.

Haggerty, S.M., Batzer, D.P., and Jackson, C.R. 2004. Macroinvertebrate response to logging in coastal headwater streams of Washington, U.S.A. Can. J. Fish. Aquat. Sci. **61**: 529–537.

Harding, J.S., Benfield, E.F., Bolstad, P.V., Helfman, G.S., and Jones, E.B.D.I. 1998. Stream biodiversity: the ghost of land use past. Proc. Natl. Acad. Sci. **95**: 14 843 – 14 847.

Hedger, R.D., Dodson, J.J., Bergeron, N.E., and Caron, F. 2005. Habitat selection by juvenile Atlantic salmon: the interaction between physical habitat and abundance. J. Fish Biol. **67**: 1054–1071.

Hicks, B.J., Hall, J.D., Bisson, P.A., and Sedell, J.R. 1991. Responses of salmonids to habitat changes. *In* Influences of forest and rangeland management on salmonid fishes and their habitats. *Edited by* W.R. Meehan. Am. Fish. Soc. Spec. Publ. No. 19, Bethesda, Maryland. pp. 483–518.

Jones, M.L., and Stockwell, J.D. 1995. A rapid procedure for the enumeration of salmonine populations in streams. N. Am. J. Fish. Manag. **15**: 551–562.

Kiffney, P.M., Richardson, J.S., and Bull, J.P. 2003. Responses of periphyton and insects to experimental manipulation of riparian buffer width along forest streams. J. Appl. Ecol. **40**: 1060–1076.

Kiffney, P.M., Richardson, J.S., and Bull, J.P. 2004. Establishing light as a causal mechanism structuring stream communities in response to experimental manipulation of riparian buffer width. J. North Am. Benthol. Soc. **23**: 542–555.

Kruse, C.G., Hubert, W.A., and Rahel, F.J. 1998. Single-pass electrofishing predicts trout abundance in mountain streams with sparse habitat. N. Am. J. Fish. Manag. **18**: 940–946.

Lamontagne, S., Carignan, R., D'Arcy, P., Prairie, Y.T., and Paré, D. 2000. Element export in runoff from eastern Canadian Boreal Shield drainage basins following forest harvesting and wildfires. Can. J. Fish. Aquat. Sci. **57**(S2): 118–128.

Latterell, J.J., Naiman, R.J., Fransen, B.R., and Bisson, P.A. 2003. Physical constraints on trout (*Oncorhynchus* spp.) distribution in the Cascade Mountains: a comparison of logged and unlogged streams. Can. J. Fish. Aquat. Sci. **60**: 1007–1017.

Lawler, J.J., and Edwards, T.C., Jr. 2002. Landscape patterns as habitat predictors: building and testing models for cavity-nesting birds in the Uinta Mountains of Utah, USA. Landsc. Ecol. **17**: 233–245.

Lobón-Cerviá, J., and Utrilla, C.G. 1993. A simple model to determine stream trout (*Salmo trutta* L.) densities based on one removal with electrofishing. Fish. Res. **15**: 369–378.

Macdonald, J.S., Beaudry, P.G., MacIsaac, E.A., and Herunter, H.E. 2003. The effects of forest harvesting and best management practices on streamflow and suspended sediment concentrations during snowmelt in headwater streams in sub-boreal forests of British Columbia, Canada. Can. J. For. Res. **33**: 1397–1407.

Madej, M.A. 1982. Sediment transport and channel changes in an aggrading stream in the Puget Lowland, Washington. *In* Sediment budgets and routing in forested drainage basins. *Edited by* F.J. Swanson, R.J. Janda, T. Dunne, and D.N. Swanston. US Department of Agriculture (USDA) Forest Service Technical Rep. No. PNW-141, Pacific Northwest Forest and Range Experimental Station, Portland, Oreg. pp. 97–109.

Martin, C.W., Hornbeck, J.W., Likens, G.E., and Buso, D.C. 2000. Impacts of intensive harvesting on hydrology and nutrient dynamics of northern hardwood forests. Can. J. Fish. Aquat. Sci. **57**(S2): 19–29.

Muñoz, J., and Felicísimo, A.M. 2004. Comparison of statistical methods commonly used in predictive modelling. J. Veg. Sci. **15**: 285–292.

Murphy, M.L., and Hall, J.D. 1981. Varied effects of clear-cut logging on predators and their habitat in small streams of the Cascade Mountains, Oregon. Can. J. Fish. Aquat. Sci. **38**: 137–145.

Murphy, M.L., Heifetz, J., Johnson, S.W., Koski, K.V., and Thedinga, J.F. 1986. Effects of clear-cut logging with and without buffer strips on juvenile salmonids in Alaskan streams. Can. J. Fish. Aquat. Sci. **43**: 1521–1533.

Quinn, T.P., and Peterson, N.P. 1996. The influence of habitat complexity and fish size on over-winter survival and growth of individually marked juvenile coho salmon (*Oncorhynchus kisutch*) in Big Beef Creek, Washington. Can. J. Fish. Aquat. Sci. **53**: 1555–1564.

© 2007 NRC Canada

Rowe, D.K., Smith, J., Quinn, J., and Boothroyd, I. 2002. Effects of logging with and without riparian strips on fish species abundance, mean size, and the structure of native fish assemblages in Coromandel, New Zealand, streams. N.Z. J. Mar. Freshw. Res. **36**: 67–79.

Rutherford, D.A., Echelle, A.A., and Maughan, O.E. 1992. Drainage-wide effects of timber harvesting on the structure of stream fish assemblages in southeastern Oklahoma. Trans. Am. Fish. Soc. **121**: 716–728.

Swank, W.T., Vose, J.M., and Elliott, K.J. 2001. Long-term hydrologic and water quality responses following commercial clearcutting of mixed hardwoods on a southern Appalachian catchment. For. Ecol. Manag. **143**: 163–178.

Swanston, D.N. 1991. Natural processes. *In* Influences of forest and rangeland management on salmonid fishes and their habitats. *Edited by* W.R. Meehan. American Fisheries Society Special Publication No. 19, Bethesda, Maryland. pp. 139–179.

Turgeon, K., and Rodríguez, M.A. 2005. Predicting microhabitat selection in juvenile Atlantic salmon *Salmo salar* by the use of logistic regression and classification trees. Freshw. Biol. **50**: 539–551.

Wiens, J.A. 1989. Spatial scaling in ecology. Funct. Ecol. **3**: 385–397.

© 2007 NRC Canada

Chinook salmon (*Oncorhynchus tshawytscha*) — seabird covariation off central California and possible forecasting applications

Jennifer E. Roth, Kyra L. Mills, and William J. Sydeman

Abstract: We evaluated covariation between Chinook salmon (*Oncorhynchus tshawytscha*) abundance and seabird breeding success in central California, USA, and compared potential forecasts to predictive models based on jack (2-year-old male) returns in the previous year. Stepwise regression models based on seabird breeding success in the previous year were comparable to or stronger than jack-based models. Including seabird breeding success in the current year improved the strength of the relationships. Combined approaches that included seabird and jack data further improved the models in some cases. The relationships based on seabird breeding success remained relatively strong over both shorter (1990–2004) and longer (1976–2004) time periods. Regression models based on multivariate seabird or combined seabird–jack indices were not as strong as stepwise regression models. Our results indicate that there is significant covariation in the responses of salmon and seabirds to variability in ocean conditions and that seabird data may offer an alternate way of forecasting salmon abundance in central California.

Résumé : Nous évaluons la covariation entre l'abondance des saumons chinook (*Oncorhynchus tshawytscha*) et le succès de la reproduction des oiseaux marins en Californie centrale, É.-U., et nous comparons les prédictions potentielles ainsi générées à celles des modèles prédictifs basés sur les retours des madeleineaux (mâles de 2 ans) de l'année précédente. Les modèles de régression pas à pas basés sur le succès de la reproduction des oiseaux marins durant l'année précédente sont aussi ou plus puissants que les modèles basés sur les madeleineaux. L'inclusion du succès de la reproduction des oiseaux marins de l'année courante augmente la force des relations. Dans certains cas, l'inclusion conjointe des données sur les oiseaux de mer et les madeleineaux améliore encore plus les modèles. Les relations basées sur le succès de la reproduction des oiseaux marins demeurent relativement fortes tant sur des périodes de temps plus courtes (1990–2004) que plus longues (1976–2004). Les modèles de régression basés sur des indices multidimensionnels relatifs aux oiseaux de mer ou conjointement aux oiseaux de mer et aux madeleineaux n'ont pas autant de puissance que les modèles de régression pas à pas. Nos résultats indiquent qu'il existe une covariation significative dans les réactions des saumons et des oiseaux marins à la variabilité des conditions océaniques et que les données sur les oiseaux marins peuvent servir d'outil de rechange pour prédire l'abondance des saumons en Californie centrale.

[Traduit par la Rédaction]

Introduction

The influences of ocean conditions and prey abundance on salmon populations are difficult to assess. Many researchers have focused on this issue and have demonstrated links between ocean conditions and salmon abundance and various life history and demographic characteristics. Studies of the effects of large-scale variation in ocean conditions have found that salmon catch and growth rates in the California Current System (CCS) are lower during El Niño – Southern Oscillation (ENSO) events and warm "regimes" that are characterized by lower ocean productivity (Mantua et al. 1997; Botsford and Lawrence 2002; Wells et al. 2006). Regional ocean conditions such as strong upwelling, low sea surface temperature, strong wind mixing, weak stratification of the water column, and low coastal sea level are indicative of productive ocean conditions and have a positive effect on salmon survival (Cole 2000; Koslow et al. 2002; Scheuerell and Williams 2005). Additionally, Peterson and Schwing (2003) show that the relative abundance of "cold-water" copepods has a positive influence on salmon survival.

Similar relationships have been found between seabird breeding parameters and variations in ocean conditions and prey availability. Ainley et al. (1995) demonstrated that breeding success is linked to large-scale climate indices. Many studies have documented short-term reductions in breeding success or complete breeding failure during severe warm-water events (Hodder and Graybill 1985; Wilson 1991; Sydeman et al. 2006) and longer-term declines in breeding success during warm "regimes" (Sydeman et al. 2001). On

Received 11 September 2006. Accepted 19 April 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 4 August 2007. J19530

J.E. Roth,[1] K.L. Mills,[2] and W.J. Sydeman. PRBO Conservation Science, 3820 Cypress Drive No. 11, Petaluma, CA 94954, USA.

[1]Corresponding author (e-mail: jroth@prbo.org).
[2]Present address: Red Rock Lakes National Wildlife Refuge, 27820 Southside Centennial Road, Lima, MT 59739, USA.

 doi:10.1139/F07-081 © 2007 NRC Canada

regional scales, seabirds exhibit earlier breeding and increased breeding success in years of stronger upwelling, lower sea surface temperature, and greater prey availability (Gaston and Smith 2001; Abraham and Sydeman 2004; Roth et al. 2005). These findings suggest that salmon and seabirds respond similarly to variation in ocean conditions.

One of the proposed mechanisms for the responses of both salmon and seabirds to temporal environmental variation is a "bottom-up" (nutrient-driven) effect of ocean productivity on prey availability (Sydeman et al. 2001; Koslow et al. 2002; Logerwell et al. 2003). There is considerable overlap in the diets of salmon and seabirds, and they may be responding similarly to fluctuations in a common prey base. Chinook salmon (*Oncorhynchus tshawytscha*) forage on a wide variety of marine invertebrates and fish, including euphausiids, copepods, crab larvae, squid, northern anchovy (*Engraulis mordax*), Pacific herring (*Clupea pallasii*), and rockfish (*Sebastes* spp.) in northern and central California (Merkel 1957; Hunt et al. 1999). In addition, Central Valley Chinook salmon enter and exit San Francisco Bay and may have considerable spatial overlap with seabirds breeding just offshore on Southeast Farallon Island (37°42′N, 123°00′W) and foraging in the surrounding area.

Seabirds breeding on the island include Cassin's auklets (*Ptychoramphus aleuticus*), common murres (*Uria aalge*), Brandt's cormorants (*Phalacrocorax penicillatus*), pelagic cormorants (*Phalacrocorax pelagicus*), and pigeon guillemots (*Cepphus columba*). These seabirds also forage on a wide variety of marine invertebrates and fish. Cassin's auklets are planktivores that primarily forage on euphausiids (Ainley et al. 1996*a*; Sydeman et al. 1997, 2001). Common murres forage on a variety of prey, including euphausiids, rockfish, market squid (*Loligo opalescens*), northern anchovy, Pacific whiting (*Merluccius productus*), and shiner surfperch (*Cymatogaster aggregata*; Ainley et al. 1996*b*; Sydeman et al. 1997, 2001). Brandt's cormorants are piscivores that forage on rockfish, flatfish (Bothidae and Pleuronectidae), plainfin midshipman (*Porichthys notatus*), and northern anchovy, whereas pelagic cormorants and pigeon guillemots forage on rockfish and sculpins (Cottidae; Ainley et al. 1990; Sydeman et al. 1997, 2001).

Based on the similarities outlined above, we propose that these seabirds and Chinook salmon are "trophic equivalents" and that seabirds can be used as indicators of the marine foraging conditions experienced by salmon during the ocean phase of their life cycle. To test this hypothesis, we evaluate covariation between Chinook salmon abundance and seabird breeding success in central California over two time scales, 1990–2004 and 1976–2004. Specifically, we evaluate the relationships between salmon abundance and seabird breeding success in the previous year, seabird breeding success averaged over 2- to 4-year periods (corresponding to the ocean phase of a typical salmon cohort), and multivariate indices (derived using principal components analysis) of seabird breeding success over multiple time scales. We compare models based on covariation with predictive models based on jack (2-year-old male) returns in the previous year. We repeat these procedures with jack returns in the previous year included in the models to evaluate the effectiveness of combined seabird–jack models in forecasting salmon abundance.

Materials and methods

Salmon data

We used estimates of the Central Valley index (CVI) and fall escapement in our analyses. The CVI is a measure of overall abundance that includes the ocean harvest south of Point Arena and total escapement (fish that "escape" the fishery and return to spawn) into the Central Valley. The Pacific Fishery Management Council (PFMC) currently uses linear regression to predict the CVI from the number of jack returns in the previous year (PFMC 2006). We also used fall escapement as a measure of abundance to assess whether relationships were stronger (or weaker) when we focused on only one of the salmon runs (groups of fish that undergo spawning migrations at the same time) in a given year. We chose the fall run because it currently makes up approximately 90% of the total escapement into the Central Valley (PFMC 2006). Ocean harvest for the fall run could not be partitioned out from overall harvest and was not included in the models.

Seabird data

We used estimates of Cassin's auklet, common murre, Brandt's cormorant, pelagic cormorant, and pigeon guillemot breeding success on Southeast Farallon Island in our analyses. We calculated mean breeding success (chicks per pair) for each species in each year based on data collected at individual nest sites (see Ainley and Boekelheide (1990) for a detailed description of collection methodology). We created additional variables for each seabird species that were comprised of running averages of breeding success for the current year (year x) and up to three previous years (years x and $x - 1$; years x, $x - 1$, and $x - 2$; years x, $x - 1$, $x - 2$, and $x - 3$). We used these running means as indices of ocean conditions over multiple time scales during the ocean phase of the Chinook salmon life cycle.

Statistical analyses

We evaluated the relationships between salmon abundance and seabird breeding success over two time periods. First, we assessed the relationships between the CVI and seabird breeding success for 1990–2004. We chose this time period because it corresponds to the time period used by the PFMC to forecast salmon abundance. Second, we repeated the analyses for 1976–2004 to evaluate the relationships over a longer time period and a greater range of environmental conditions. We used the same methods to evaluate relationships between fall escapement and seabird breeding success. Fall escapement numbers for 1976–2004 were log-transformed to normalize the data. We considered relationships significant at $p < 0.05$, though we focused primarily on coefficient of determination (r^2) values to evaluate covariation in regression models. We conducted all analyses in Stata 8.0 (Stata Corporation 2003).

We tested all time series data for positive autocorrelation using Durbin–Watson statistics at $p < 0.05$ (Kutner et al. 2005). The test statistic is compared with lower (d_l) and upper (d_u) bounds. Values less than d_l indicate that the data are serially correlated, values greater than d_u indicate that the data are independent, and values in between d_l and d_u indicate that the results are inconclusive. There was no evidence of autocorrelation in the seabird data for any species from 1990–

© 2007 NRC Canada

2004 (d ranged from 1.59 to 2.26 compared with d_u = 1.36) or for most species from 1976–2004 (d ranged from 2.17 to 2.35 compared with d_u = 1.48). There may be slight autocorrelation in the Cassin's auklet data from 1976–2004 (d = 1.47 compared with d_u = 1.48), though the test was inconclusive. There was no evidence of autocorrelation in the CVI for either time period (1990–2004, d = 2.01 compared with d_u = 1.36; 1976–2004, d = 1.67 compared with d_u = 1.48). There may be slight autocorrelation in fall escapement numbers from 1990–2004 (d = 1.31 compared with d_u = 1.36). There was clear evidence of autocorrelation in fall escapement numbers from 1976–2004 (d = 0.83 compared with d_l = 1.34). Autocorrelation may have increased the chance of making type I errors when we assessed relationships between fall escapement and seabird breeding success. We accepted this possibility because our primary interests were in assessing covariation between salmon abundance and seabird breeding success and evaluating potential forecasting applications (i.e., we were not testing the hypothesis that seabird breeding success directly influences salmon abundance and we were not overly concerned with p values). Furthermore, autocorrelation may be present, in part, because the species are responding to longer-term changes in ocean conditions. We, therefore, did not want to remove effects that might be evidence that the two taxa are responding similarly to those changes.

We used linear regression to evaluate covariation between salmon abundance (CVI or fall escapement) and the number of jack returns in the previous year and seabird breeding success over multiple and varying time scales (Legendre and Legendre 1998). Initially, we conducted preliminary analyses to assess univariate relationships between each seabird species and salmon abundance. We evaluated both linear and quadratic terms at $p < 0.01$ because of the large number of comparisons. Common murre breeding success was excluded from further analyses because there were no significant univariate relationships with salmon abundance. We used forward stepwise regression to determine which of the remaining seabird species were related to salmon numbers. We included Cassin's auklet, Brandt's cormorant, pelagic cormorant, and pigeon guillemot breeding success as independent variables in each stepwise regression. We conducted a stepwise regression using breeding success in the previous year (year $x - 1$). We focused on the previous year to evaluate the possibility that data collected one year in advance of the salmon returns – fishing season could be used for forecasting. We conducted separate stepwise regressions for seabird breeding success in the current year (year x) and averaged over 2 (years x and $x - 1$), 3 (years x, $x - 1$, and $x - 2$), and 4 (years x, $x - 1$, $x - 2$, and $x - 3$) years to examine the strength of the relationships over multiple time scales during the ocean phase of a typical salmon year class (i.e., potential cohort). We repeated the stepwise regressions with jack returns in the previous year included as an independent variable to assess whether combined seabird–jack approaches could strengthen predictive relationships.

We used principal components analysis to assess whether multivariate seabird indices provided stronger relationships to salmon abundance than stepwise regression analyses (Legendre and Legendre 1998). As above, we developed indices for seabird breeding success in the previous year, in the current year, and averaged over 2, 3, and 4 years. We retained principal components with eigenvalues > 1. We used linear regression to assess covariation between salmon abundance and the retained seabird principal components. We repeated this procedure with jack returns in the previous year to assess whether combined seabird–jack multivariate indices were stronger than those based on seabird data alone.

We chose the best models from the analyses described above based on adjusted r^2 values that allowed us to compare models with varying numbers of observations and variables (Kutner et al. 2005). The method tends to favor models with more variables than other model-selection procedures (e.g., Akaike's information criterion). We accepted this possibility because we were interested in the forecasting capability of the models (i.e., we did not want to penalize them too heavily for including a second or third variable). We used a cross-validation procedure to assess the ability of those models to predict salmon numbers (Kleinbaum et al. 1988). We excluded each year in turn from each of the analyses, estimated regression equations using the remaining data, and predicted the values for the excluded years. We used linear regression to compare predicted and actual values for each model. We used the resulting r^2 value (cross-validation r^2) to assess how well the model would perform when applied to new data. We also calculated the average percent deviation between predicted and actual values.

Results

The predictive relationship between the CVI and jack returns in the previous year was strong for 1990–2004 (adjusted r^2 = 0.61) and weak for 1976–2004 (adjusted r^2 = 0.16; Table 1). The best seabird model based on data from the previous year included Cassin's auklet breeding success for both time periods. The seabird model was comparable to the jack model for 1990–2004 (adjusted r^2 = 0.60) and was an improvement over the jack model for 1976–2004 (adjusted r^2 = 0.33), though it is noteworthy that the predictive value was lower over the longer time period for both jacks and seabirds. Combining the seabird and jack data did not improve the relationships for 1990–2004. The combined model was a substantial improvement over other models for 1976–2004, explaining an additional 44% of the variation in the CVI compared with the jack model. This model included jack returns and Cassin's auklet, Brandt's cormorant, and pigeon guillemot breeding success.

Fall escapement was weakly related to fall jack returns in the previous year for 1990–2004 (adjusted r^2 = 0.26), and the relationship was not significant for 1976–2004 (adjusted r^2 = 0.07; Table 1). The best seabird models were much better and included Cassin's auklet breeding success for both time periods. These models explained 25%–30% more variation in fall escapement than the jack models and were significant over both shorter and longer time periods. Combining the seabird and jack data did not improve the relationships for either time period.

The best models for predicting the CVI based on multiyear averages of seabird data included a 2-year average (years x and $x - 1$) of Brandt's cormorant breeding success

© 2007 NRC Canada

Table 1. Comparisons of models relating salmon abundance to the number of jacks returning to spawn, seabird breeding success, and combinations of jack and seabird data in the previous year.

	Central Valley index		Fall escapement	
	1990–2004	1976–2004	1990–2004	1976–2004
Jack models				
Jacks	+	+	+	
Model results				
Adjusted r^2	0.61	0.16	0.26	0.07
p value	<0.001	<0.05	<0.05	>0.05
Seabird models				
CAAU	+	+	+	+
BRAC				
PIGU				
PECO				
Model results				
Adjusted r^2	0.60	0.33	0.56	0.32
p value	<0.001	<0.001	<0.001	<0.001
Combined models				
Jacks	+	+		
CAAU		+	+	+
BRAC		+		
PIGU		–		
PECO				
Model results				
Adjusted r^2	0.61	0.60	0.56	0.32
p value	<0.001	<0.001	<0.001	<0.001

Note: Variables included in all of the models were lagged by 1 year. Jack models were based on simple linear regressions. Seabird and combined models were based on stepwise regressions. Significant variables are denoted with a plus or minus depending on the sign of the regression coefficient. CAAU, Cassin's auklet; BRAC, Brandt's cormorant; PIGU, pigeon guillemot; PECO, pelagic cormorant.

for 1990–2004 (adjusted r^2 = 0.60) and 2-year averages of Cassin's auklet and Brandt's cormorant breeding success for 1976–2004 (adjusted r^2 = 0.64; Table 2). These seabird models were comparable to the jack model for 1990–2004 (adjusted r^2 = 0.61) and an improvement over the jack model for 1976–2004 (adjusted r^2 = 0.16). The best combined models strengthened the relationships for both time periods, explaining an additional 19% and 54% of the variation in the CVI compared with the jack models.

The best seabird model for predicting fall escapement was one including a 2-year average of Cassin's auklet breeding success for 1990–2004 (adjusted r^2 = 0.78; Table 3). Three-year averages of Cassin's auklet and Brandt's cormorant breeding success were important in predicting fall escapement for 1976–2004 (adjusted r^2 = 0.66). The seabird models explained an additional 52% and 59% of the variation in fall escapement compared with the jack models and were better than the combined models for both time periods.

The best models for predicting salmon numbers based on multivariate seabird indices were generally not as strong as the comparable stepwise regression models. For indices based on data from the previous year, only the combined index for predicting the CVI from 1990–2004 had a higher adjusted r^2 value (adjusted r^2 = 0.72; Table 4) than the comparable stepwise regression model (adjusted r^2 = 0.61). The models for predicting the CVI and fall escapement based on multiyear averages of seabird breeding success had lower adjusted r^2 values than the comparable stepwise regression models for both the seabird and combined indices (Tables 5 and 6).

The seabird–jack index based on multiyear averages of seabird breeding success had the least shrinkage (i.e., the lowest drop in the r^2 value) when its ability to predict the CVI for 1990–2004 was tested (Table 7). The combined model based on multiyear averages had the lowest average percent deviation between predicted and actual values for that time period (Fig. 1). For 1976–2004, the seabird index based on multiyear averages of seabird breeding success had the least shrinkage, whereas the best combined seabird–jack model based on multiyear averages had the lowest average percent deviation between predicted and actual values. The best seabird models for predicting fall escapement based on multiyear averages had the least shrinkage and the lowest average percent deviation between predicted and actual values for 1990–2004. For 1976–2004, the seabird model and seabird index based on multiyear averages had the least shrinkage and the seabird model had the lowest average percent deviation between predicted and actual values.

Discussion

Our results indicate that there is covariation in the responses of seabirds and Chinook salmon to variability in

© 2007 NRC Canada

Table 2. Comparisons of models relating the Central Valley index to seabird breeding success and a combination of seabird breeding success and the number of jacks returning to spawn in the previous year.

	Current year (year x)		Two-year average (years x and $x-1$)		Three-year average (years x, $x-1$, and $x-2$)		Four-year average (years x, $x-1$, $x-2$, and $x-3$)	
	1990–2004	1976–2004	1990–2004	1976–2004	1990–2004	1976–2004	1990–2004	1976–2004
Seabird models								
CAAU	+	+		+	+			
BRAC			+	+		+	+	+
PIGU								
PECO								
Model results								
Adjusted r^2	0.26	0.28	**0.60**	**0.64**	0.51	0.53	0.37	0.53
p value	<0.05	<0.01	**<0.001**	**<0.001**	<0.01	<0.001	<0.01	<0.001
Combined models								
Jacks	+	+	+	+	+	+	+	+
CAAU		+		+				
BRAC			+	+	+	+		+
PIGU								
PECO								
Model results								
Adjusted r^2	0.61	0.38	**0.80**	**0.70**	0.70	0.67	0.61	0.58
p value	<0.001	0.001	**<0.001**	**<0.001**	<0.001	<0.001	<0.001	<0.001

Note: Values in bold indicate models with the highest adjusted r^2 values for each time period. Stepwise regression models were based on breeding success in the current year and averages of the current year and up to three previous years. All regression coefficients were positive.

Table 3. Comparisons of models relating fall escapement to seabird breeding success and a combination of seabird breeding success and the number of fall jacks returning to spawn in the previous year.

	Current year (year x)		Two-year average (years x and $x-1$)		Three-year average (years x, $x-1$, and $x-2$)		Four-year average (years x, $x-1$, $x-2$, and $x-3$)	
	1990–2004	1976–2004	1990–2004	1976–2004	1990–2004	1976–2004	1990–2004	1976–2004
Seabird models								
CAAU	+	+	+	+	+	+	+	+
BRAC						+		+
PIGU								
PECO								
Model results								
Adjusted r^2	0.57	0.49	**0.78**	0.63	0.71	**0.66**	0.64	0.58
p value	<0.001	<0.001	**<0.001**	<0.001	<0.001	**<0.001**	<0.001	<0.001
Combined models								
Jacks								
CAAU	+	+	+	+	+	+	+	+
BRAC						+		+
PIGU								
PECO								
Model results								
Adjusted r^2	0.57	0.49	**0.78**	0.63	0.71	**0.66**	0.64	0.58
p value	<0.001	<0.001	**<0.001**	<0.001	<0.001	**<0.001**	<0.001	<0.001

Note: Values in bold indicate models with the highest adjusted r^2 values for each time period. Stepwise regression models were based on breeding success in the current year and averages of the current year and up to three previous years. All regression coefficients were positive.

ocean conditions and that seabird data may be useful in a forecasting context. Others have discussed the use of seabirds as indicators of forage fish abundance and zooplankton availability (Monaghan et al. 1989; Cairns 1992; Abraham and Sydeman 2006). However, most previous studies have focused on "nowcasting" of fish abundance rather than the predictive capabilities of the relationships. Velarde et al. (2004) provide an exception to this, with a focus on predictive models of Pacific sardine (*Sardinops caeruleus*) catch in Baja California that incorporate seabird diet and breeding

© 2007 NRC Canada

Table 4. Comparisons of models relating salmon abundance to seabird indices (principal components) and combined seabird–jack indices using data from the previous year.

	Central Valley index		Fall escapement	
	1990–2004	1976–2004	1990–2004	1976–2004
Seabird models[a]				
Factor 1	+	+	+	+
Model results				
Adjusted r^2	0.27	0.23	0.42	0.15
p value	<0.05	<0.01	<0.01	<0.05
Combined models[b]				
Factor 1	+	+	+	+
Factor 2	+	NA	NS	NA
Model results				
Adjusted r^2	0.72	0.27	0.48	0.17
p value	<0.001	<0.01	<0.01	<0.05

Note: Factors were derived from principal components analyses. NA, the second factor was not retained for the regression analyses shown above (eigenvalue < 1); NS, the factor was not significant and was subsequently dropped from the regression analyses. All regression coefficients were positive.

[a]Principal components analysis included data on CAAU, BRAC, PIGU, and PECO.

[b]Principal components analysis included seabird breeding success for all species listed above and jacks in the previous year.

Table 5. Comparisons of models relating the Central Valley index to seabird indices (principal components) and combined seabird–jack indices.

	Current year (year x)		Two-year average (years x and $x-1$)		Three-year average (years x, $x-1$, and $x-2$)		Four-year average (years x, $x-1$, $x-2$, and $x-3$)	
	1990–2004	1976–2004	1990–2004	1976–2004	1990–2004	1976–2004	1990–2004	1976–2004
Seabird models[a]								
Factor 1	+	+	+	+	+	+	+	+
Model results								
Adjusted r^2	0.34	0.22	**0.49**	**0.46**	0.38	0.31	0.24	0.24
p value	<0.01	<0.01	**<0.01**	**<0.001**	<0.01	<0.001	<0.05	<0.01
Combined models[b]								
Factor 1	+	+	+	+	+	+	+	+
Factor 2	NA	+	NA	NA	NA	+	+	NA
Model results								
Adjusted r^2	0.46	0.39	**0.63**	0.50	0.50	**0.56**	0.58	0.28
p value	<0.01	<0.001	**<0.001**	<0.001	<0.01	**<0.001**	<0.01	<0.01

Note: Values in bold indicate models with the highest adjusted r^2 values for each time period. Factors were derived from principal components analyses. NA, the second factor was not retained for regression analyses. All regression coefficients were positive.

[a]Based on seabird breeding success in the current year and averages of the current year and up to three previous years.

[b]Based on jacks in the previous year and seabird breeding success in the current year and averages of the current year and up to three previous years.

parameters. Velarde et al.'s (2004) effort is based, in part, on the direct consumption of sardines by the seabirds. Our effort is unique in that our models are based on an apparent similarity in the responses of specific seabirds and Chinook salmon that is presumably due to the similarity in their diets. Indeed, the one species that consumes juvenile Chinook salmon, common murre, was dropped from our modeling effort early on. The covariation we found offers support for the idea that seabirds and salmon are "trophic equivalents" and are both being influenced by "bottom-up" (nutrient-driven) processes. In the California Current System, ecosystem productivity is influenced by two main factors: (*i*) the strength of the California Current that brings nutrient-rich waters into the region from the north and (*ii*) seasonal upwelling that brings nutrients to the surface (Chelton et al. 1982; Schwing et al. 2000). If the seabirds in this study directly influenced Chinook salmon in a "top-down" fashion, we would have found negative correlations. Although we cannot rule out the possibility that other top predators are affecting both seabirds and salmon, we suggest that a "bottom-up" mechanism is most compelling.

We found significant relationships between salmon abundance and seabird breeding success in the previous year and seabird breeding success averaged over multiple years. Models including averages of the current (year x) and 1 (year $x-1$) or 2 (years $x-1$ and $x-2$) previous years were stronger than those based only on the previous year. Though we were not able to analyze the data by salmon year class (cohort), our results suggest that ocean conditions in the later phases of the salmon life cycle at sea are important.

© 2007 NRC Canada

Table 6. Comparisons of models relating fall escapement to seabird indices (principal components) and combined seabird–jack indices.

	Current year (year x)		Two-year average (years x and $x - 1$)		Three-year average (years x, $x - 1$, and $x - 2$)		Four-year average (years x, $x - 1$, $x - 2$, and $x - 3$)	
	1990–2004	1976–2004	1990–2004	1976–2004	1990–2004	1976–2004	1990–2004	1976–2004
Seabird models[a]								
Factor 1	+	+	+	+	+	+	+	+
Model results								
Adjusted r^2	0.52	0.40	**0.73**	**0.54**	0.70	0.48	0.54	0.35
p value	0.001	<0.001	**<0.001**	**<0.001**	<0.001	<0.001	0.001	<0.001
Combined models[b]								
Factor 1	+	+	+	+	+	+	+	+
Factor 2	NA	NS	NS	NA	NS	+	NS	+
Model results								
Adjusted r^2	0.59	0.42	**0.76**	0.56	0.73	**0.61**	0.60	0.46
p value	<0.001	<0.001	**<0.001**	<0.001	<0.001	**<0.001**	<0.001	<0.001

Note: Values in bold indicate models with the highest adjusted r^2 values for each time period. Factors were derived from principal components analyses. NA, the second factor was not retained for regression analyses; NS, the second factor was not significant and was subsequently dropped from the regression analyses. All regression coefficients were positive.

[a]Based on seabird breeding success in the current year and averages of the current year and up to three previous years.

[b]Based on jacks in the previous year and seabird breeding success in the current year and averages of the current year and up to three previous years.

Our results are consistent with studies that have documented influences of ocean conditions in both the year of ocean entry and the year of return on coho salmon (*Oncorhynchus kisutch*) growth and survival in the northern California Current System (Botsford and Lawrence 2002; Koslow et al. 2002; Wells et al. 2006). Relationships between Chinook salmon and ocean conditions are generally more complex as a result of wide variation in age-at-return, but ocean conditions in the year of return appear to be important (Botsford and Lawrence 2002; Wells et al. 2006). Scheuerell and Williams (2005) developed predictive models of Snake River Chinook salmon survival, finding that including upwelling in the year of return substantially improved the predictive capability of their models. There is, however, a trade-off between the quality of the prediction and data availability in our models. Seabird data from the current year are not available before the start of the fishing season and would, therefore, be difficult to include in management decisions.

Patterns across time scales

The strength of the relationships between salmon abundance and both seabird breeding success and jack returns was generally weaker over the longer time period, 1976–2004. This result is not surprising given the greater range of marine conditions that salmon and seabirds are likely to experience in longer versus shorter time periods and the associated difficulty in capturing that variation (Sydeman et al. 2001; Chavez et al. 2003). However, relationships between salmon abundance and seabird breeding success and combined seabird–jack data remained relatively strong in comparison to models based only on jacks, which were poor for the longer time series. These results are encouraging because relationships between salmon survival and physical indices of ocean conditions have tended to break down over time. For instance, Lawson (1997) found that an established relationship between coho salmon survival and upwelling was not significant during a later period of weaker upwelling. Miller and Sydeman (2004) noted a similar breakdown between upwelling and rockfish productivity. Consequently, researchers have included multiple physical variables in their analyses in order to better capture the variation in the ocean environment (Koslow et al. 2002; Logerwell et al. 2003). Our results suggest that including seabird data in fisheries models may offer an additional method for capturing the important biological and physical variation driving fish population dynamics.

Univariate versus multivariate approaches

The models based on individual variables (stepwise regressions) were generally stronger than those based on multivariate indices (principal components). We developed the multivariate indices to evaluate the possibility that a single seabird index or a single combined seabird–jack index could be developed that would simplify analytical considerations and interpretations. However, the seabirds included in the indices have different sensitivities to variation in ocean conditions because of differences in reproductive strategies and trophic ecology (Ainley et al. 1995; Sydeman et al. 2001), and combining them into a single index appears to obscure some of the covariation evident in single-species analyses. The fact that not all species were included in stepwise regression models supports this interpretation.

Interestingly, Cassin's auklet breeding success was included in most of the stepwise regression models. Cassin's auklets may be especially good indicators of ocean conditions for salmon because they feed extensively on euphausiids off central California (Ainley et al. 1996*a*; Abraham and Sydeman 2004, 2006). Euphausiids probably represent the lowest trophic level directly exploited by seabirds and salmon and may be one of the most responsive prey items to changes in ocean characteristics. Brandt's cormorant breeding success was also included in many of the stepwise regression models based on multiyear averages of seabird breeding success. Brandt's cormorants feed on a wide variety of fish species in neritic habitats (Sydeman et al. 1997) and may be good indicators of average ocean conditions over longer time periods be-

© 2007 NRC Canada

Table 7. Comparisons of the predictive capability of the best models from each analysis, including r^2 values from the original models, cross-validation r^2 values, shrinkage, and the average percent deviation between predicted and actual values.

	r^2 [a]	Cross-validation r^2 [b]	Shrinkage[c]	Average percent deviation ± SD
Central Valley index (1990–2004)				
Previous year				
Jack model	0.64	0.54	0.10	5±25
Seabird model	0.62	0.48	0.14	7±39
Combined model	NA	NA	NA	NA
Seabird index	0.32	0.17	0.15	14±58
Seabird–jack index	0.76	0.59	0.17	5±30
Multiyear averages				
Seabird model	0.63	0.49	0.14	5±40
Combined model	0.83	0.73	0.10	1±22
Seabird index	0.53	0.42	0.11	9±38
Seabird–jack index	0.65	0.56	0.09	6±31
Central Valley index (1976–2004)				
Previous year				
Jack model	0.19	0.09	0.10	12±40
Seabird model	0.36	0.27	0.09	9±37
Combined model	0.66	0.50	0.16	5±34
Seabird index	0.26	0.17	0.09	11±43
Seabird–jack index	0.29	0.21	0.08	10±42
Multiyear averages				
Seabird model	0.66	0.58	0.08	4±25
Combined model	0.73	0.63	0.10	2±24
Seabird index	0.48	0.42	0.06	7±30
Seabird–jack index	0.59	0.51	0.08	5±26
Fall escapement (1990–2004)				
Previous year				
Jack model	0.31	0.18	0.13	29±59
Seabird model	0.59	0.46	0.13	17±77
Combined model	NA	NA	NA	NA
Seabird index	0.46	0.30	0.16	33±118
Seabird–jack index	0.51	0.38	0.13	26±98
Multiyear averages				
Seabird model	0.80	0.72	0.08	2±45
Combined model	NA	NA	NA	NA
Seabird index	0.75	0.65	0.10	11±63
Seabird–jack index	0.78	0.68	0.10	6±48
Fall escapement (1976–2004)				
Previous year				
Jack model	0.10	0.02	0.08	16±56
Seabird model	0.34	0.25	0.09	10±52
Combined model	NA	NA	NA	NA
Seabird index	0.18	0.07	0.11	14±64
Seabird–jack index	0.20	0.10	0.10	14±62
Multiyear averages				
Seabird model	0.68	0.61	0.07	5±33
Combined model	NA	NA	NA	NA
Seabird index	0.56	0.49	0.07	7±40
Seabird–jack index	0.64	0.53	0.11	8±44

Note: The best models from simple and stepwise regressions for each estimate of salmon abundance and each time period were tested using leave-one-out cross-validation. SD, standard deviation; NA, either jack returns in the previous year or seabird breeding success did not enter the combined stepwise regression models.

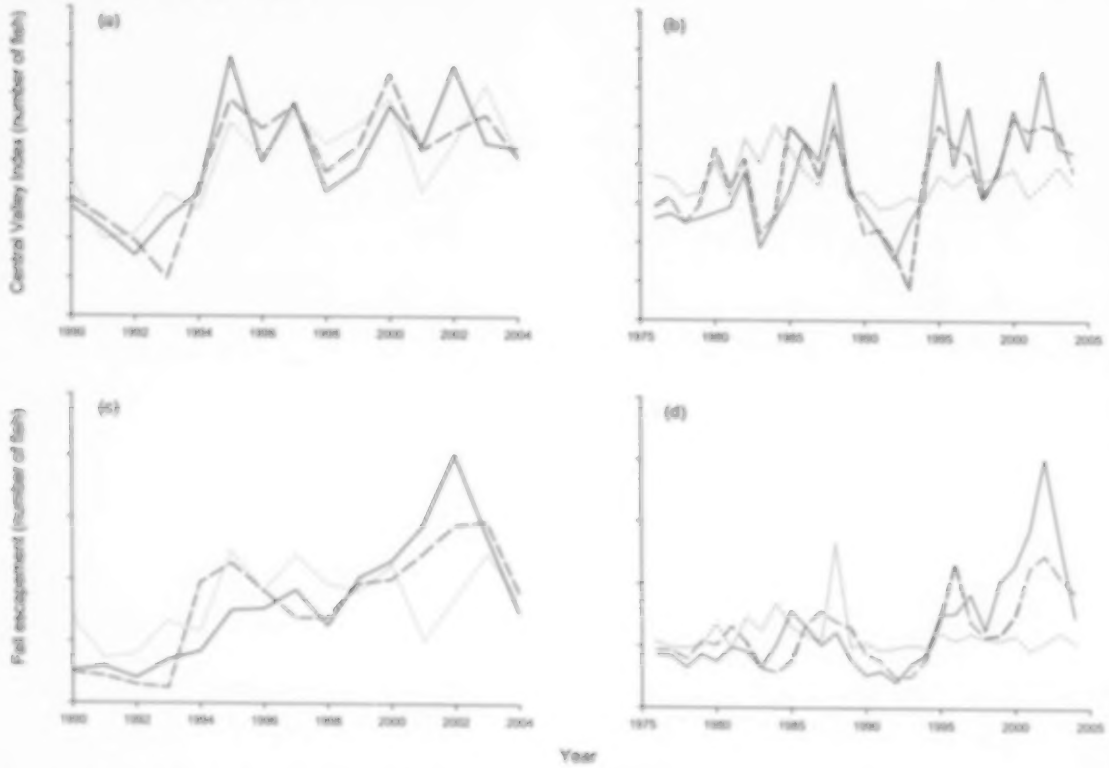
[a] Note that these are original r^2 values and do not match adjusted r^2 values shown in other tables.

[b] Cross-validation r^2 values reflect the relationships between predicted and actual salmon abundances.

[c] Difference between the r^2 values of the original models and the cross-validation r^2 values that reflects the drop in the amount of variation explained by the models when they were tested.

© 2007 NRC Canada

Fig. 1. Comparisons of actual and predicted salmon numbers for (*a*) the Central Valley index from 1990–2004, (*b*) the Central Valley index from 1976–2004, (*c*) fall escapement from 1990–2004, and (*d*) fall escapement from 1976–2004. The solid line represents actual abundance estimates in all graphs, the dotted line represents predicted abundances based on jacks in the previous year in all graphs, and the broken line represents predicted abundances based on combined seabird–jack models in (*a*) and (*b*) and predicted abundances based on seabird models in (*c*) and (*d*).



cause of their reliance on fish populations that exhibit complex age structures and, therefore, respond more slowly than zooplankton to ocean changes. Spatial distribution of foraging seabirds also may explain some of our results. Brandt's cormorants forage over the continental shelf, whereas Cassin's auklets forage primarily over outer continental shelf and slope waters (Ainley et al. 1996*a*; Sydeman et al. 1997; Yen et al. 2004). Relationships between Brandt's cormorants and salmon at longer time lags may reflect greater use of nearshore waters by both cormorants and younger salmon, though little is known about age-specific use of marine habitats by salmon.

Potential applications

Botsford et al. (1997) argued that an ecosystem approach to fisheries management is important to improving forecasts of fish abundance and maintaining sustainable fisheries. In addition, they stressed the importance of including both biological and physical parameters in predictive models. Seabird breeding success on the Farallon Islands is related to both biological and physical factors (Ainley et al. 1995; Abraham and Sydeman 2004) and integrates variation in the marine environment from egg laying (March) through chick rearing (August) each year. Our results indicate that this integrative parameter can be used to monitor marine ecological conditions important to other upper-trophic-level organisms such as predatory fishes and may have a role in ecological forecasting. Seabird breeding success is relatively inexpensive to measure and has been calibrated against ocean conditions and prey abundance in some locations. Including seabird data in fisheries models may, therefore, enhance understanding of the factors affecting fish populations and may be valuable in a management context in regions where data on commercially important fish species and seabirds is available.

Acknowledgements

The Farallon National Wildlife Refuge is managed by the US Fish and Wildlife Service (USFWS). Support for seabird studies on the Farallones was provided by a variety of agencies and organizations, including USFWS, California Department of Fish and Game, Packard Foundation, Gordon and Betty Moore Foundation, Bradford Foundation, ExxonMobile Corporation, and the members and donors of PRBO Conservation Science. Salmon data were provided courtesy of the Pacific Fishery Management Council. Resources Law Group, Moore Family Foundation, and the National Fish and Wildlife Foundation provided financial support for this analysis. Allen Grover, Michael Mohr, and David Welch shared their knowledge of salmon research, forecasting, and management. Morten Frederiksen and an anonymous reviewer provided helpful

© 2007 NRC Canada

comments on the manuscript. This is PRBO contribution No. 1327.

References

Abraham, C.L., and Sydeman, W.J. 2004. Ocean climate, euphausiids and auklet nesting: inter-annual trends and variation in phenology, diet and growth of a planktivorous seabird, *Ptychoramphus aleuticus*. Mar. Ecol. Prog. Ser. **274**: 235–250.

Abraham, C.L., and Sydeman, W.J. 2006. Prey-switching by Cassin's auklet *Ptychoramphus aleuticus* reveals seasonal climate-related cycles of *Euphausia pacifica* and *Thysanoessa spinifera*. Mar. Ecol. Prog. Ser. **313**: 271–283.

Ainley, D.G., and Boekelheide, R.J. 1990. Seabirds of the Farallon Islands: ecology, dynamics, and structure of an upwelling-system community. Stanford University Press, Stanford, California.

Ainley, D.G., Strong, C.S., Penniman, T.M., and Boekelheide, R.J. 1990. The feeding ecology of Farallon seabirds. *In* Seabirds of the Farallon Islands: ecology, dynamics, and structure of an upwelling-system community. *Edited by* D.G. Ainley and R.J. Boekelheide. Stanford University Press, Stanford, California. pp. 51–127.

Ainley, D.G., Sydeman, W.J., and Norton, J. 1995. Upper trophic level predators indicate interannual negative and positive anomalies in the California Current food web. Mar. Ecol. Prog. Ser. **118**: 69–79.

Ainley, D.G., Spear, L.B., and Allen, S.G. 1996*a*. Variation in the diet of Cassin's Auklet reveals spatial, seasonal, and decadal occurrence patterns of euphausiids off California, USA. Mar. Ecol. Prog. Ser. **137**: 1–10.

Ainley, D.G., Spear, L.B., Allen, S.G., and Ribic, C.A. 1996*b*. Temporal and spatial patterns in the diet of the Common Murre in California waters. Condor, **98**: 691–705.

Botsford, L.W., and Lawrence, C.A. 2002. Patterns of co-variability among California Current chinook salmon, coho salmon, Dungeness crab, and physical oceanographic conditions. Progr. Oceanogr. **53**: 283–305.

Botsford, L.W., Castilla, J.C., and Peterson, C.H. 1997. The management of fisheries and marine ecosystems. Science (Washington, D.C.), **277**: 509–515.

Cairns, D.K. 1992. Bridging the gap between ornithology and fisheries science: use of seabird data in stock assessment models. Condor, **94**: 811–824.

Chavez, F.P., Ryan, J., Lluch-Cota, S.E., and Ñiquen, C.M. 2003. From anchovies to sardines and back: multidecadal change in the Pacific Ocean. Science (Washington, D.C.), **299**: 217–221.

Chelton, D.B., Bernal, P.A., and McGowan, J.A. 1982. Large-scale interannual physical and biological interaction in the California Current. J. Mar. Res. **40**: 1095–1125.

Cole, J. 2000. Coastal sea surface temperature and coho salmon production off the north-west United States. Fish. Oceanogr. **9**: 1–16.

Gaston, A.J., and Smith, J.L. 2001. Changes in oceanographic conditions off northern British Columbia (1983–1999) and the reproduction of a marine bird, the Ancient Murrelet (*Synthliboramphus antiquus*). Can. J. Zool. **79**: 1735–1742.

Hodder, J., and Graybill, M.R. 1985. Reproduction and survival of seabirds in Oregon during the 1982–1983 El Niño. Condor, **87**: 535–541.

Hunt, S.L., Mulligan, T.J., and Komori, K. 1999. Oceanic feeding habits of chinook salmon, *Oncorhynchus tshawytscha*, off Northern California. Fish. Bull. **97**: 717–721.

Kleinbaum, D.G., Kupper, L.L., and Muller, K.E. 1988. Applied regression analysis and other multivariable methods. 2nd ed. PWS–KENT Publishing Company, Boston, Massachusetts.

Koslow, J.A., Hobday, A.J., and Boehlert, G.W. 2002. Climate variability and marine survival of coho salmon (*Oncorhynchus kisutch*) in the Oregon production area. Fish. Oceanogr. **11**(2): 65–77.

Kutner, M.H., Nachtsheim, C.J., Neter, J., and Li, W. 2005. Applied linear statistical models. McGraw-Hill/Irwin, Boston, Massachusetts.

Lawson, P.W. 1997. Interannual variability in growth and survival of Chinook and coho salmon. *In* Estuarine and Ocean Survival of Northeastern Pacific Salmon: Proceedings of the Workshop. *Edited by* R.L. Emmett and M.H. Schiewe. United States Department of Commerce, National Oceanic and Atmospheric Administration Technical Memo. No. NMFS-NWFSC-29. pp. 81–91.

Legendre, P., and Legendre, L. 1998. Numerical ecology. 2nd English ed. Elsevier, Amsterdam, the Netherlands.

Logerwell, E.A., Mantua, N., Lawson, P.W., Francis, R.C., and Agostini, V.N. 2003. Tracking environmental processes in the coastal zone for understanding and predicting Oregon coho (*Oncorhynchus kisutch*) marine survival. Fish. Oceanogr. **12**(6): 554–568.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., and Francis, R.C. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bull. Am. Meteorol. Soc. **78**(6): 1069–1079.

Merkel, T.J. 1957. Food habits of the king salmon (*Oncorhynchus tshawytscha* (Walbaum)) in the vicinity of San Francisco, California. Calif. Fish Game, **43**(4): 249–270.

Miller, A.K., and Sydeman, W.J. 2004. Rockfish response to low-frequency ocean climate change as revealed by the diet of a marine bird over multiple time scales. Mar. Ecol. Prog. Ser. **281**: 207–216.

Monaghan, P., Uttley, J.D., and Okill, J.D. 1989. Terns and sandeels: seabirds as indicators of changes in marine fish populations. J. Fish Biol. **35**(Suppl. A): 339–340.

Pacific Fishery Management Council. 2006. Preseason report I: stock abundance analysis for 2006 ocean salmon fisheries. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 200, Portland, OR 97220-1384, USA. Available from http://www.pcouncil.org/salmon/salpre.html [accessed 28 February 2007].

Peterson, W.T., and Schwing, F.B. 2003. A new climate regime in Northeast Pacific ecosystems. Geophys. Res. Lett. **30**, 1896. doi:10.1029/2003GL017528.

Roth, J.E., Sydeman, W.J., and Martin, P.L. 2005. Xantus's murrelet breeding relative to prey abundance and oceanographic conditions in the Southern California Bight. Mar. Ornithol. **33**(2): 115–121.

Scheuerell, M.D., and Williams, J.G. 2005. Forecasting climate-induced changes in the survival of Snake River spring/summer Chinook salmon (*Oncorhynchus tshawytscha*). Fish. Oceanogr. **14**(6): 448–457.

Schwing, F.B., Moore, C.S., Ralston, S., and Sakuma, K.M. 2000. Record coastal upwelling in the California Current in 1999. Calif. Coop. Ocean. Fish. Investig. Rep. **41**: 148–160.

Stata Corporation. 2003. Stata user's guide. Release 8. Stata Press, College Station, Texas.

Sydeman, W.J., Hobson, K.A., Pyle, P., and McLaren, E.B. 1997. Trophic relationships among seabirds in central California: combined stable isotope and conventional dietary approach. Condor, **99**: 327–336.

© 2007 NRC Canada

Sydeman, W.J., Hester, M.M., Thayer, J.A., Gress, F., Martin, P., and Buffa, J. 2001. Climate change, reproductive performance and diet composition of marine birds in the Southern California Current System, 1969–1997. Prog. Oceanogr. **49**: 309–329.

Sydeman, W.J., Bradley, R.W., Warzybok, P., Abraham, C.L., Jahncke, J., Hyrenbach, K.D., Kousky, V., Hipfner, J.M., and Ohman, M.D. 2006. Planktivorous auklet *Ptychoramphus aleuticus* responses to ocean climate, 2005: unusual atmospheric blocking? Geophys. Res. Lett. **33**: L22S09. doi:10.1029/2006GL026736.

Velarde, E., Ezcurra, E., Cisneros-Mata, M.A., and Lavin, M.F. 2004. Seabird ecology, El Niño anomalies, and prediction of sardine fisheries in the Gulf of California. Ecol. Appl. **14**(2): 607–615.

Wells, B.K., Grimes, C.B., Field, J.C., and Reiss, C.S. 2006. Covariation between the average lengths of mature coho (*Oncorhynchus kisutch*) and Chinook salmon (*O. tshawytscha*) and the ocean environment. Fish. Oceanogr. **15**(1): 67–79.

Wilson, U.W. 1991. Responses of three seabird species to El Niño events and other warm episodes on the Washington Coast, 1979–1990. Condor, **93**: 853–858.

Yen, P.P.W., Sydeman, W.J., and Hyrenbach, K.D. 2004. Marine bird and cetacean associations with bathymetric habitats and shallow-water topographies: implications for trophic transfer and conservation. J. Mar. Syst. **50**: 79–99.

© 2007 NRC Canada

Geometric morphometric analysis of fish scales for identifying genera, species, and local populations within the Mugilidae

Ana L. Ibañez, Ian G. Cowx, and Paul O'Higgins

Abstract: Geometric morphometric methods (GMMs) were used to determine if scale morphology can discriminate between genera, species, geographic variants, and stocks of mullet (Mugilidae). GMMs were used because they allow standard multivariate analyses while preserving information about scale shape, which is important in making biological interpretations of results. The method was tested on ctenoid scales from mullets collected from different areas of the Gulf of Mexico and Aegean Sea. Scales were submitted to generalised procrustes analysis, followed by principal components analysis of resulting shape coordinates. Principal component scores were submitted to cross-validated discriminant analysis to determine the efficacy of scale landmarks in discriminating by taxon and population. Fish scale form was least effective in discriminating populations from nearby areas, better when populations are more geographically dispersed, and best between species and genera. Scale form variations reflected previous genetic studies that differentiated congeneric *Mugil cephalus* and *Mugil curema*, which are distinct from other Mugilidae. The method is nondestructive, quick, and less costly than genetic analysis, thus allowing many individuals to be screened.

Résumé : Nous utilisons des méthodes morphométriques géométriques (« GMMs ») pour vérifier si la morphologie de l'écaille permet de distinguer les genres, les espèces, les variantes géographiques et les stocks chez les mulets (Mugilidae). Nous avons choisi les GMMs parce qu'elles permettent les analyses multidimensionnelles habituelles, tout en conservant l'information sur la forme des écailles qui est importante dans l'interprétation biologique des résultats. Nous testons la méthode sur les écailles cténoïdes de mulets récoltés dans diverses régions du golfe du Mexique et de la mer Égée. Les écailles ont été soumises à l'analyse généralisée du logiciel procrustes, puis les coordonnées de forme obtenues à l'analyse en composantes principales. Les positions des composantes principales ont été soumises à une analyse discriminante à validation croisée afin de déterminer l'efficacité des repères sur l'écaille pour distinguer les taxons et les populations. La forme de l'écaille de poisson permet de discriminer moins bien entre les populations de régions adjacentes, mieux entre les populations plus éloignées géographiquement et le mieux entre les espèces et les genres. Les variations de forme des écailles reflètent les résultats d'études génétiques antérieures qui distinguent les congénères *Mugil cephalus* et *Mugil curema* qui sont différents des autres Mugilidae. La méthode ne requiert pas la destruction de l'échantillon, elle est rapide et moins coûteuse qu'une analyse génétique, ce qui permet de trier un nombre élevé d'individus.

[Traduit par la Rédaction]

Introduction

Species identification and population discrimination are important in the conservation of biodiversity, natural resources, and fisheries management. It is also often necessary to identify individual specimens correctly to population when investigating biological traits such as growth, mortality, fecundity, trophic relationships, parasite relationships, and historical and paleontological events.

Traditionally, identification of species membership has relied on morphological and meristic characteristics (Casselman et al. 1981; Ihssen et al. 1981; Cadrin 2000), but these methods often encounter problems with reliability (e.g., Jerry and Cairns 1998; Swain and Foote 1999; Murta 2000). More recently, genetic methods have been used (e.g., Hutchinson et al. 2001; Mariani et al. 2005), but they are costly and not readily available in the field. Alternatively, Fourier analysis of fish scales has been used to discriminate between stocks (e.g., Jarvis et al. 1978; Richards and Esteves 1997; Poulet et al. 2005). More recently, Watkinson and Gillis (2005) used wavelet analysis as an alternative to Fourier analysis to improve discrimination. Alternative methods of shape analysis, based on landmark data (and geometric morphometric methods (GMMs)), have also found wide ap-

Received 31 May 2006. Accepted 13 April 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 4 August 2007. J19345

A.L. Ibañez[1,2] **and I.G. Cowx.** Hull International Fisheries Institute, Hull HU6 7RX, UK.
P. O'Higgins. Hull York Medical School and Department of Biology, The University of York, Heslington, York YO10 5DD, UK.

[1]Corresponding author (e-mail: ana@xanum.uam.mx).
[2]Present address: Universidad Autónoma Metropolitana-Iztapalapa, Departamento de Hidrobiología, Av San Rafael Atlixco 186, Col. Vicentina, México, D.F. 09340, México.

 doi:10.1139/F07-075 © 2007 NRC Canada

plicability in biology (Chen et al. 2005; Crews and Hedin 2006; Recasens et al. 2006) because of the natural links between homologies and measurements (the landmark coordinates; cf. Fourier and wavelet approaches), the statistical properties of the resulting shape spaces (Dryden and Mardia 1998), and good statistical power. These GMMs preserve geometry throughout the analysis and, like Fourier and wavelet analyses, provide graphical visualisations of the statistical findings that can aid biological interpretation (Walker 1996; O'Higgins 2000). Different measurement methods (e.g., Fourier, outline curvatures, wavelets) result in varying assessments of the differences between specimens to the extent that the choice of method is arbitrary and the resulting distance matrices (the elements of which are the relative distances between specimens) will also be arbitrary. The landmark-based methods differ fundamentally from others in that they utilise biological equivalences between structures, matching or "homologous" landmarks, and to this extent, the resulting metrics are founded in biology. Having said this, in practical terms, the Fourier and other "landmark-free" approaches work well and offer the immediate advantage of capturing the subtleties of outline detail. GMMs have advanced to incorporate these outline features (e.g., Bookstein et al. 2005) and it is expected that these approaches will, in time, offer all of the advantages of outline measurement methods while incorporating all available homology information. However, the extent to which outline variations between landmarks are informative depends greatly on the features of the structure being measured, the adequacy of sampling of form that can be achieved by landmarks alone, and the question at hand. In the case of fish scales, a reasonable representation of form is achieved using very few landmarks (as we will show further on), and there are definite advantages to achieving a concise description of form by using a few variables because the resulting multivariate spaces are readily visualised and "noise" is minimised relative to "signal" in discriminant analysis.

According to Thomson (1981), the family Mugilidae contains 14 genera and 64 species that live in coastal waters and estuaries at all latitudes, except polar regions. The external morphology of this family is highly conservative, which contributes to uncertainty with regard to evolutionary relationships at both generic and specific levels (Crosetti et al. 1994). Mullets are euryhaline species that can penetrate lagoons and estuaries, migrating back to the sea to spawn. They play an important part in small-scale coastal fisheries in several regions of the world. These fisheries constitute one of the 10 most important coastal fisheries in Mexico as a result of their catch volume, which exceeds 13 000 tonnes annually. Most of the catch takes place in the northern Gulf of Mexico in the Madre and Tamiahua lagoons. In Greek coastal lagoons, mullets represent the 11th target of small-scale fisheries, accounting for 2.3% of total catches (Tzanatos et al. 2005).

The aim of this study was to evaluate whether the landmark-based, geometric, morphometric approach to describing fish scale morphology is useful in discriminating between species, populations, and geographic variants. This methodology was assessed using a case study based on fish species of the family Mugilidae collected from different localities from the Gulf of Mexico and Aegean Sea. The null hypothesis being tested was that differences as described by landmarks in fish scale morphology between species and populations of mullet are insufficient to allow identification by genus, species, geographic variants, and local populations.

Materials and methods

Fish scale collection/age and size range of fishes

Scales of *Mugil cephalus* and *Mugil curema* were collected in Tamiahua Lagoon, Mexico, between April 1991 and March 1992. Further fish scales from these species were collected in December 2005 from the Mexico City central fish market: *M. cephalus* specimens came from Madre Lagoon and *M. curema* specimens came from Tecolutla River. All Mexican localities are near the Gulf of Mexico (Table 1; Figs. 1*a*, 1*b*).

Samples from the Aegean Sea were obtained from fish caught on the east coast of Greece in January 2006 (Figs. 2*a*, 2*b*): *M. cephalus* scales were acquired from Monolimni Lagoon on the River Evros and Agiasma Lagoon; *Chelon labrosus* (Risso 1827) scales were acquired from the Vassova, Eratino, and Agiasma lagoons (Nestos River); and *Liza saliens* (Risso 1810) were acquired from the Eratino Lagoon. Fish were measured to the nearest millimetre (mm) in situ, and scales were removed from the shoulder region in front of the 1st dorsal fin ray above the lateral line and stored dry in paper envelopes. In the laboratory, scales were cleaned with soft soap and tap water, dried, and examined using a microfiche projector. A digital image was taken from the flat screen of the projector. Only one scale per fish was used for the analysis (Table 1) as we are attempting to evaluate a rapid stock discrimination methodology that does not require repetitive sampling of individuals. All specimens were adults except for the sample of *M. cephalus* from the Tamiahua Lagoon, which contained 40% juveniles. In this paper, "area" refers to a large water body such as the Gulf of Mexico or Aegean Sea, and "local population" refers to specimens collected within a small geographic area such as a lagoon in the Gulf of Mexico or on the Greek coast.

Morphometrics

Preliminary visual assessment was used to identify potential landmarks on the scales. The landmarks chosen are located on key features of the scale that are common to all scales of the species or variants examined. This ensures that in subsequent interpretation of results, variations in particular landmarks can be related back to shared features of form (contrast with Fourier analysis where no such relationship exists except for the start point for the series; see Introduction). For mullet scales, the following landmarks were considered appropriate (Fig. 3): landmarks 1 and 3 are the ventro- and dorso-lateral tips of the anterior portion of the scale, respectively; landmark 2 is in the center of the anterior edge of the scale, landmarks 4 and 6 are at the boundary between the anterior area with circuli and the posterior area covered by ctenі (spine-like ornamentations), respectively; landmark 5 is the focus of the scale; and landmark 7 is positioned at the tip of the posterior portion of the scale.

© 2007 NRC Canada

Table 1. Size of sample and site of collection of fish scales.

Species	Area	Collection site	Group code	No. of fish	Fork length range (cm)	Sample collection period
Mugil cephalus (*Mce*)	GM	Ma	*Mce*-GMMa	33	27.0–31.0*	December 2005
		Ta	*Mce*-GMTa	29	22.5–35.6*	April 1991 to March 1992
	AS	Ag	*Mce*-ASAg	10	28.5–32.5	January 2006
		Mo	*Mce*-ASMo	40	24.8–47.2	January 2006
Mugil curema (*Mcu*)	GM	Ta	*Mcu*-GMTa	30	19.5–31.0*	April 1991 to March 1992
		Te	*Mcu*-GMTe	13	27.5–36.0*	December 2005
Chelon labrosus (*Cla*)	AS	Va	*Cla*-ASVa	10	23.3–29.6	January 2006
		Er	*Cla*-ASEr	27	20.0–28.6	January 2006
		Ag	*Cla*-ASAg	10	24.0–27.0	January 2006
Liza saliens (*Lsa*)	AS	Er	*Lsa*-ASEr	5	22.2–24.1*	January 2006

Note: Areas: AS, Aegean Sea; GM, Gulf of Mexico. Collection sites: Ag, Agiasma Lagoon; Er, Eratino Lagoon; Ma, Madre Lagoon; Mo, Monolimni Lagoon; Ta, Tamiahua Lagoon; Te, Tecolutla River; Va, Vassova Lagoon.

*Specimens were originally measured with total length, which was converted to fork length according to the equation given by Thompson et al. (1991).

Fig. 1. Sampling locations in the Gulf of Mexico: (*a*) Madre Lagoon; (*b*) Tamiahua Lagoon and Tecolutla River.

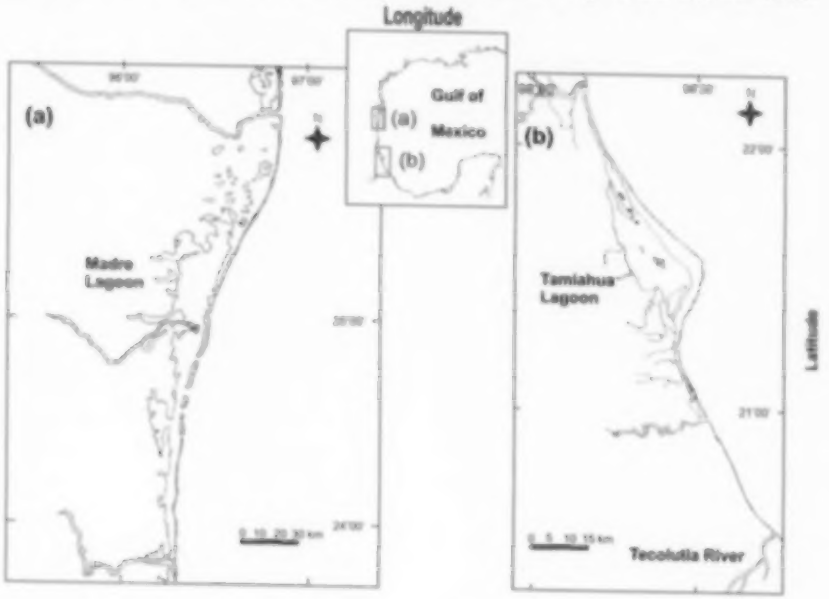
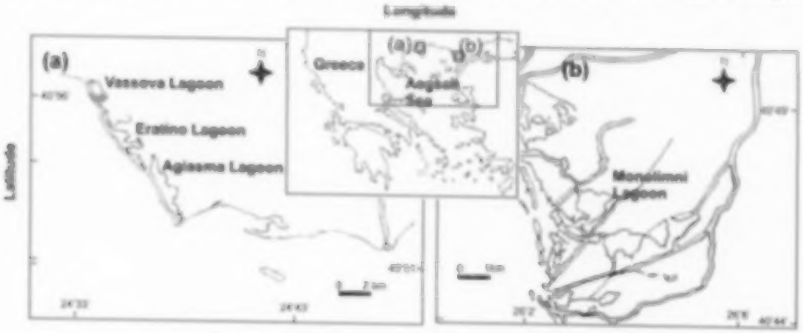
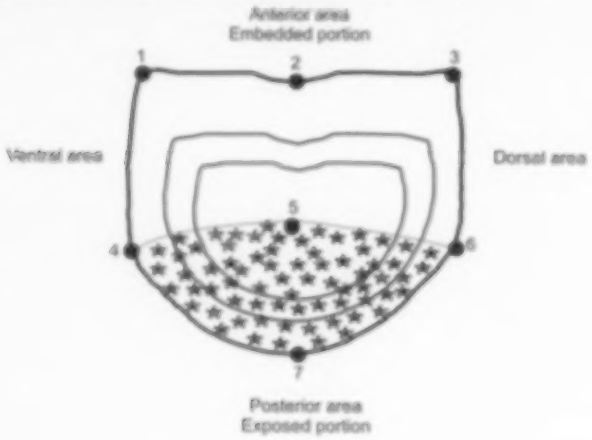


Fig. 2. Sampling locations in the Aegean Sea: (*a*) Agiasma, Eratino, and Vassova lagoons; (*b*) Monolimni Lagoon.



© 2007 NRC Canada

Fig. 3. Landmark definitions used in the fish scales.



The configurations of landmark coordinates were scaled, translated, and rotated using generalised procrustes analysis (GPA). They were then submitted to tangent projection (Dryden and Mardia 1993) and subsequently to principal components analysis (PCA) (Dryden and Mardia 1993; Kent 1994). The aim was to examine the potential for differences in shape to allow classification of unknown specimens, and so the scores of specimens on all non-zero principal components (PCs) were submitted to discriminant analysis (SPSS ver. 13.0) to compute generalised Mahalanobis distances and discriminant functions and to assess the efficacy of the latter in classification. This was carried out using cross-validation in which multiple repeated analyses were carried out leaving out one individual in the construction of the discriminant function before classifying this individual according to the function. This avoids overestimating the efficacy of discriminant functions by using them to classify specimens employed in their construction. Percentage correct classification rates were recorded. The aspects of shape variability represented by the discriminant functions were estimated by multivariate regression of PC scores on discriminant function scores. The PC scores predicted by the extremes of each discriminant function were then used to reconstruct the expected shape of landmark configurations with those particular scores by adding to the mean tangent coordinates the products of these PC scores and the eigenvectors for those PCs before projecting back from the tangent to the configuration space (O'Higgins et al. 2001). Finally, the differences in shape between the mean and the shapes represented by the extremes of discriminant functions of interest were visualised using transformation grids computed using thin-plate splines (TPS) (Bookstein 1989; Marcus et al. 1996; Dryden and Mardia 1998).

Sequential analyses were carried out to assess the extent to which scale shape differences, as described by landmarks, allow identification by genus, species, geographic variant, and local population.

To assess the quality of discrimination by genus relative to that by species, GPA of the whole sample (207 fish scales) was carried out, and the PC scores were submitted to two discriminant analyses: one in which the scales were labeled by genus and the other, by species. Differences in fish scale shape between geographic variants were assessed by GPA/PCA and subsequent discriminant analysis of the *M. cephalus* specimens (112 fish scales), 62 from the Gulf of Mexico and 50 from the Aegean Sea. Differences in fish scale shape between species and geographic variants were assessed by GPA/PCA and subsequent discriminant analysis of the PC scores of specimens of the genus *Mugil* (155 fish scales). Classifications of specimens were carried out by species (*M. cephalus* and *M. curema*) and geographic variant, Gulf of Mexico and Aegean Sea. Finally, the differences between local populations were assessed using the same methods applied to *M. cephalus* from the Gulf of Mexico (62 fish scales: 33 for Madre Lagoon and 29 for Tamiahua Lagoon) and Greece (50 specimens: 10 for Agiasma and 40 for Monolimni Lagoon), *M. curema* from the Gulf of Mexico (43 specimens: 30 for Tamiahua Lagoon and 13 for Tecolutla River), and *C. labrosus* from the three Greek lagoons (47 specimens: 10 fish scales for Vassova, 27 for Eratino, and 10 for Agiasma Lagoon).

Results

The first canonical discriminant function for the 207 individuals classified by genus explained 90.7% of the total between-genus variance, whereas the second accounted for 9.3% (Wilks' lambda = 0.560, $p < 0.001$). The discriminant analysis correctly classified 74.4% of the fish scales to genus, whereas the cross-validated analysis correctly classified 71.0% of the fish scales (Table 2). The best classification rate, 73.5%, was for *Mugil* compared with 63.8% and 60.0%, respectively, for *Chelon* and *Liza*; most misclassifications occurred between the latter two genera (Table 2).

Three groupings were defined in the plot of the first two discriminant functions from the analysis of the 207 individuals. The centroids and specimens of *M. cephalus* and *M. curema* were separated principally on the first discriminant function, whereas those of *C. labrosus* and *L. saliens* occupied a central position with considerable overlap of specimens and lower scores than *M. cephalus* and *M. curema* on the second function (Fig. 4). The first discriminant function explained 79.2% of the total variance, whereas the second accounted for 17.6% (Wilks' lambda = 0.252, $p < 0.001$). The discriminant analysis correctly classified 78.7% of the original grouped cases, whereas cross-validation correctly classified 73.9% (Table 3). *Mugil cephalus* has the best classification rate, with 83.9% correctly classified after cross-validation. Misclassifications were common between *M. curema* and *C. labrosus* and between *L. saliens*, *M. cephalus*, and *C. labrosus* (Table 3).

The shape variation represented by the first discriminant function of Fig. 4 is shown (Fig. 5), computed by regressing shape on that function. The leftmost grid (Fig. 5*a*) represents the mean shape warped to a score of –0.4, the central grid (Fig. 5*b*) represents the mean with a score of 0, and the rightmost grid (Fig. 5*c*) represents a score of 4.0. Scales from *C. labrosus*, *L. saliens*, and *M. curema* (which lie to the negative pole of the first discriminant function and so have shapes that differ from the mean as represented in Fig. 5*a*) were characterized by a relatively shorter distance

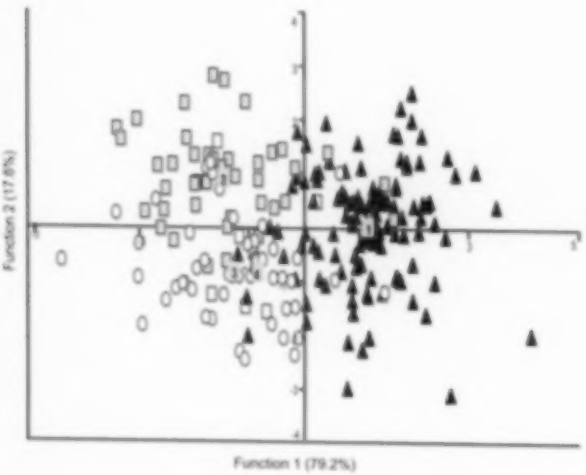
© 2007 NRC Canada

Table 2. Classification results[a] for the cross-validation testing procedure for the three genera of the Mugilidae family.

Genus	Predicted group membership			Total
	Mugil	*Chelon*	*Liza*	
Count				
Mugil	114	27	14	155
Chelon	5	30	12	47
Liza	1	1	3	5
Percent				
Mugil	73.5	17.4	9.0	100.0
Chelon	10.6	63.8	25.5	100.0
Liza	20.0	20.0	60.0	100.0

[a]Of cross-validated grouped cases, 71.0% were correctly classified.

Fig. 4. Scatter plot of the discriminant function scores from the analysis of species: solid triangles, *Mugil cephalus*; shaded squares, *M. cuvema*; open circles, *Chelon labrosus*; open squares, *Liza saliens*. Numbers over the shaded squares are the number of centroids.



between the focus and landmark 7, with a convex anterior edge, whereas this distance for *M. cephalus* is relatively larger, with the focus more central and the anterior edge concave (Fig. 5*c*). Likewise the shape variation represented by the second discriminant function visualised with a transformation grid is illustrated (Fig. 6). From left to right, each grid portrays the shape for the sample with the following discriminant function scores (DS): –3.0 (Fig. 6*a*), 0.0 (Fig. 6*b*), and 3.0 (Fig. 6*c*). The degree of shape variation appears less than in Fig. 5, in part because the transformation grids span a smaller range on the discriminant function. The key difference is in the relative location of the landmark, which becomes more posterior with respect to the remaining landmarks. The grid is distorted principally between this landmark and the focus. Thus, the first discriminant function of Fig. 4 principally relates to contrasts between groups in the relative locations of the focus and landmark 2, whereas the second relates to contrasts in

Table 3. Classification results[a] for the cross-validation testing procedure for the four mullet species.

Species	Predicted group membership				Total
	Mce	*Mcu*	*Cla*	*Lsa*	
Count					
Mce	94	4	7	7	112
Mcu	3	30	5	5	43
Cla	3	5	26	13	47
Lsa	1	0	1	3	5
Percent					
Mce	83.9	3.6	6.3	6.3	100.0
Mcu	7.0	69.8	11.6	11.6	100.0
Cla	6.4	10.6	55.3	27.7	100.0
Lsa	20.0	0.0	20.0	60.0	100.0

Note: *Mce*, *M. cephalus*; *Mcu*, *M. cuvema*; *Cla*, *Chelon labrosus*; *Lsa*, *Liza saliens*.

[a]Of cross-validated grouped cases, 73.9% were correctly classified.

Fig. 5. Shape variation along the first discriminant function of Fig. 4, visualized using a transformation grid. From left to right, each grid shows the predicted shape for DS: (*a*) –3.0, (*b*) 0.0, and (*c*) 3.0.

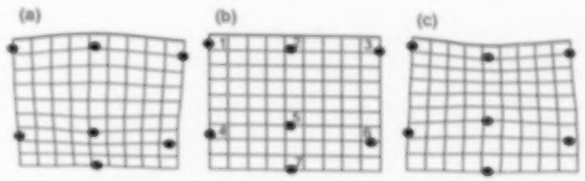
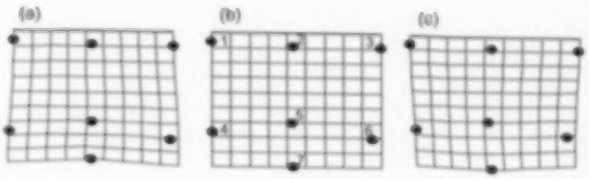


Fig. 6. Shape variation along the second discriminant function of Fig. 4, visualized using a transformation grid. From left to right, each grid shows the predicted shape for DS: (*a*) –3.0, (*b*) 0.0, and (*c*) 3.0.



landmark 7. All of these landmarks are arranged along the axis of the scale, and so it is variation in the longitudinal axis that accounts for the discrimination between groups.

The first discriminant function from the analysis of the 112 fish scales of *M. cephalus* specimens explained 100% of the total variance and was highly significant (Wilks' lambda = 0.320, $p < 0.001$). The discriminant analysis correctly classified 92.0% of the original grouped cases; whereas cross-validation resulted in 91.1% correct classification (Table 4).

Three groups were distinguished by the discriminant analysis of the *Mugil* species (155 individuals) by area (Wilks' lambda = 0.138, $p < 0.001$). The first function (71.5% of the total variance; Fig. 7) separated *M. cephalus* from the Gulf of Mexico and the Aegean Sea and both species of *Mugil*;

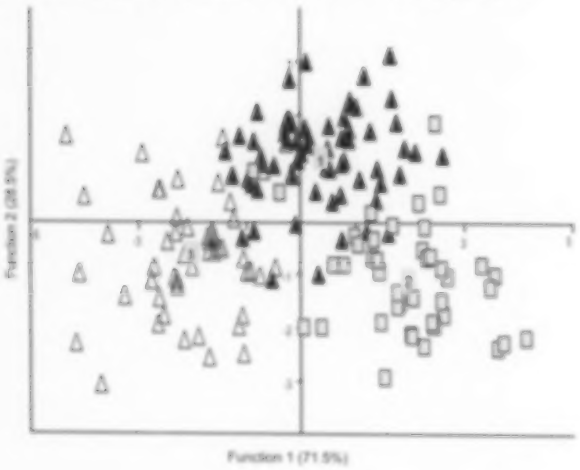
© 2007 NRC Canada

Table 4. Classification results[a] for the cross-validation testing procedure for *Mugil cephalus* in the Gulf of Mexico (GM) and Aegean Sea (AS).

	Predicted group membership		
Areas	GM	AS	Total
Count			
GM	56	6	62
AS	4	46	50
Percent			
GM	90.3	9.7	100.0
AS	8.0	92.0	100.0

[a]Of cross-validated grouped cases, 91.1% were correctly classified.

Fig. 7. Scatter plot of the discriminant scores from the analysis of the *Mugil* species: solid triangles, *M. cephalus* from the Gulf of Mexico; open triangles, *M. cephalus* from the Aegean Sea; shaded squares, *M. curema* from the Gulf of Mexico. Numbers over the shaded squares represent the number of centroids.



the second function (28.5% of the total variance) separated *M. cephalus* from the Gulf of Mexico from *M. curema* and *M. cephalus* from the Aegean Sea. The discriminant analysis correctly classified 90.3% of the original grouped cases and 89.7% when cross-validated (Table 5). *Mugil cephalus* from the Aegean Sea was classified with 92% accuracy in the original and cross-validated analyses, whereas *M. cephalus* and *M. curema* from the Gulf of Mexico were correctly classified in 87.1% and 90.7% of cases, respectively (Table 5).

The shape variation represented by the first discriminant function of Fig. 7 is visualized using transformation grids, showing a DS of –4.0 (Fig. 8*a*), the mean (Fig. 8*b*), and the score 4.35 (Fig. 8*c*). As DS I increases, the exposed portion of the scale becomes relatively narrower (between landmarks 4 and 6) and the embedded portion becomes relatively broader (between landmarks 1 and 3). Likewise, the shape variation represented by the second discriminant function is illustrated (Fig. 9). From left to right, each grid shows DS –3.0, 0, and 3.0 (Figs. 9*a*, 9*b*, 9*c*, respectively). The deformations of panels in Figs. 9*a* and 9*c* indicate that with increasing DS II, the

Table 5. Classification results[a] for the cross-validation testing procedure for the *Mugil* species (*M. cephalus*, *M. curema*) in the Gulf of Mexico (GM) and Aegean Sea (AS).

		Predicted group membership			
		M. cephalus		*M. curema*	
Species	Area	GM	AS	GM	Total
Count					
M. cephalus	GM	54	4	4	62
	AS	4	46	0	50
M. curema	GM	4	0	39	43
Percent					
M. cephalus	GM	87.1	6.5	6.5	100.0
	AS	8.0	92.0	0.0	100.0
M. curema	GM	9.3	0.0	90.7	100.0

[a]Of cross-validated grouped cases, 89.7% were correctly classified.

Fig. 8. Shape variation along the first discriminant function of Fig. 7, visualized using a transformation grid. From left to right, each grid shows the predicted shape for DS: (*a*) –4.0, (*b*) 0.0, and (*c*) 4.0.

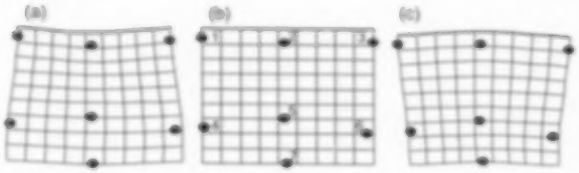
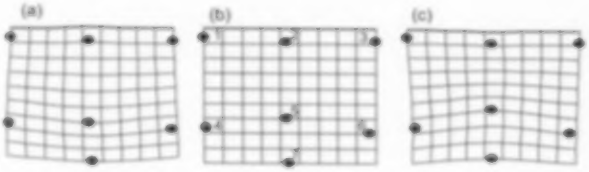


Fig. 9. Shape variation along the second discriminant function of Fig. 7, visualized using a transformation grid. From left to right, each grid shows the predicted shape for DS: (*a*) –3.0, (*b*) 0.0, and (*c*) 3.0.



focus becomes more central with a little narrowing between landmarks 4 and 6.

The cross-validated classification results from the analysis between local populations for the *Mugil* species (Table 6) identified significant discrimination in all cases ($p < 0.001$). The cross-validated results for *M. cephalus* from the Gulf of Mexico showed a 74.2% correct overall classification: 72.7% for Madre Lagoon and 75.9% for Tamiahua Lagoon (Wilks' lambda = 0.536, $p < 0.001$). The cross-validated results for the Greek lagoons correctly classified 82.0% of specimens overall: with 70.0% and 85.0% for Monolimni and Agiasma lagoons, respectively (Wilks' lambda = 0.565, $p < 0.001$). *Mugil curema* from the Gulf of Mexico was correctly classified 74.4% overall: 83.3% for Tamiahua Lagoon and 53.8% for Tecolutla River (Wilks' lambda = 0.463, $p < 0.001$).

Lower cross-validated correct classification rates were obtained for *C. labrosus* with respect to the three Greek coastal

© 2007 NRC Canada

Table 6. Classification results for the cross-validation testing procedure for pair local populations of *Mugil* (*M. cephalus* (*Mce*), *M. curema* (*Mcu*)).

		Predicted group membership						
		Mce				*Mcu*		
Species	Area	GMMa	GMTa	ASAg	ASMo	GMTa	GMTe	Total
Count								
Mce	GMMa	24	9					33
	GMTa	7	22					29
	ASAg			7	3			10
	ASMo			6	34			40
Mcu	GMTa					25	5	30
	GMTe					6	7	13
Percent								
Mce	GMMa	72.7	27.3					100.0[a]
	GMTa	24.1	75.9					100.0[a]
	ASAg			70.0	30.0			100.0[b]
	ASMo			15.0	85.0			100.0[b]
Mcu	GMTa					83.3	16.7	100.0[c]
	GMTe					46.2	53.8	100.0[c]

Note: GMMa, Gulf of Mexico, Madre Lagoon; GMTa, Gulf of Mexico, Tamiahua Lagoon; ASAg, Aegean Sea, Agiasma Lagoon; ASMo, Aegean Sea, Monolimni Lagoon; GMTe, Gulf of Mexico, Tecolutla River.

[a]Of cross-validated grouped cases, 74.2% were correctly classified

[b]Of cross-validated grouped cases, 82.0% were correctly classified.

[c]Of cross-validated grouped cases, 74.4% were correctly classified.

lagoons. The first canonical discriminant function explained 69.3% of the total variance, whereas the second function accounted for 30.7%, but overall discrimination was significant (Wilks' lambda = 0.434, $p < 0.001$). The discriminant analysis correctly classified 70.2% of the original grouped cases, whereas the cross-validated procedure correctly classified 48.9% (Table 7). *Chelon labrosus* from Eratino Lagoon gave the best classification rate, 55.6%, whereas 40% of specimens from Vassova and Agiasma lagoons were correctly classified.

Table 7. Classification results[a] for the cross-validation testing procedure for local populations of *Chelon labrosus* in Nestos River lagoons from the Aegean Sea.

	Predicted group membership			
Areas	ASVa	ASEr	ASAg	Total
Count				
ASVa	4	3	3	10
ASEr	7	15	5	27
ASAg	3	3	4	10
Percent				
ASVa	40.0	30.0	30.0	100.0
ASEr	25.9	55.6	18.5	100.0
ASAg	30.0	30.0	40.0	100.0

Note: ASVa, Aegean Sea, Vassova Lagoon; ASEr, Aegean Sea, Eratino Lagoon; ASAg, Aegean Sea, Agiasma Lagoon.

[a]Of cross-validated grouped cases, 48.9% were correctly classified.

Discussion

The analysis provides sufficient evidence to reject the null hypothesis; scale morphology is appropriate to allow discrimination by genera, species, geographic variants, and local populations.

Although the discrimination between genera and species in all cases was over 60%, there are similarities between the clusters comprising *C. labrosus* and *L. saliens*. Despite the limited sample of *L. saliens*, results were consistent with the original and cross-validated classifications in the genus and species discrimination analyses. The close resemblance of these two species has been reported based on chromosome complements and mitochondrial DNA sequence analysis (Cataudella et al. 1974; Papasotiropoulos et al. 2001, 2002). Turan et al. (2005) also suggested that the separation of *Chelon* and *Liza* into separate genera might be inappropriate, based on allozyme electrophoresis. Similarly, Guinea and Fernandez (1992) showed broad similarities between the gill rake morphology of three species of *Liza* and *C. labrosus*. However, despite *M. curema* and *C. labrosus* sharing similarities in the shape of the scales, it is possible to discriminate between them in a large proportion of cases, even with seven landmarks. This might be further improved by increasing the number of landmarks. Thus, discrete differences between species are prevalent, supporting their separate species nomenclature.

The genus *Mugil* was even better segregated, with *M. cephalus* well separated from other species. *Mugil curema* more closely resembled *C. labrosus* and *L. saliens*. These results agree with genetic studies of the Mugilidae that indicate remarkable genetic divergence in *M. cephalus* and *M. curema* compared with other members of the family (Caldara et al. 1996).

The correct classification rates based on fish scale shape differences between populations of *M. cephalus* agree with genetic analyses, which indicates extensive genetic diversity

© 2007 NRC Canada

between populations from around the world (Crosetti et al. 1994; Rossi et al. 1998; Rocha-Olivares et al. 2000). Rocha-Olivares et al. (2000) identified the existence of a single population within the northern Gulf of Mexico. In the present study, specimens from different lagoons in the Gulf were well separated, pointing to the prevalence of several stocks. Similarly, *M. cephalus* in Greek waters could be reasonably well identified, but *C. labrosus* stocks were poorly discriminated. Not all differences should be attributed solely to genetic differences; environmental influences might also influence fish morphology (West-Eberhard 1989; Shepherd 1991; Kinsey et al. 1994). Such environmental influences on morphology have led to the definition of phenotypic stocks characterized by phenotypic differences that may be exclusively environmentally induced (Swain and Foote 1999). Temporal differences in sampling might also be a source of variability. The 13- to 14-year time lag between the collection periods of the *Mugil* species from the Gulf of Mexico might have confounded geographic variation to some extent, but further, more-intensive sampling will address this issue. Allometry presents a further confounding source of shape variability that we were unable to take into account in the present study. Similar effects are seen in otoliths that undergo morphology changes with age and environmental conditions (Felix-Uraga et al. 2005; Monteiro et al. 2005). Nevertheless, the aim of this study was to evaluate a rapid assessment tool for stock discrimination, and it is expected that accounting for ontogenic variability will only improve the output.

The variable discrimination between regions and species might be explained by the life history of mullets (Ibañez and Gallardo-Cabello 2004; Ibañez and Gutiérrez-Benítez 2004). *Mugil cephalus* in the Gulf of Mexico waters spawn in depths of 40–1650 m and beyond the continental shelf 50–100 km off the coast (Arnold and Thompson 1958). They do not migrate extensively (Idyll and Sutton 1951; Funicelli et al. 1989) and return to their original coastal area after spawning (Mahmoudi 1991). The larvae remain in the coastal waters for 2 to 3 months (Ditty and Shaw 1996) where they are confined to surface currents that transport them to estuarine and coastal regions that could not be the same as those of their parents. Similar reproductive behaviour is shown by *M. cephalus*, *C. labrosus*, and *L. saliens* from the Mediterranean Sea (Koutrakis et al. 1994; Koutrakis 2004), where floating eggs and larvae of mullet are transported by sea currents (Rossi 1986). The distances between the mouths of the Madre and Tamiahua lagoons and between the Tamiahua Lagoon and the Tecolutla River are ~370 km and ~96 km, respectively; these distances seem sufficient to maintain stock integrity. The extent to which they can be identified by lagoon depends on the balance between larval migration and adult stasis. Similarly for the Aegean systems, Monolimni Lagoon belongs to the River Evros delta, which is some 100 km from Agiasma Lagoon, a part of the River Nestos delta.

By contrast, the samples of *C. labrosus* were taken from three coastal water bodies, the Vassova, Eratino, and Agiasma lagoons, which form part of the delta of the River Nestos; the mouths of the Vassova and Eratino are only 3.7 km apart, whereas the Eratino and Agiasma are separated by 5.3 km, and there is a possible narrow connection between Vassova and Eratino (Sotiris Orfanidis, National Agricultural Research Foundation Fisheries Research Institute, Kavala, Greece, personal communication, 2006). These geographical conditions go a long way to explaining the poor discrimination rates as they are likely predisposed to considerable gene flow with similar environmental conditions.

Although landmarks varied between the groups of fish in the present study, the central axes formed by landmarks 2, 5, and 7 appear most significant in the analysis of all 207 scales. The distinctiveness of the *M. cephalus* fish scales with respect to all other species is mainly due to a more centrally placed focus. Scales of the genus *Mugil* are generally relatively longer and narrower than those of other taxa within the Mugilidae. Such scale shape differences may well reflect adaptations to different functions and swimming characteristics, but there is a paucity of information with relation to these species in the literature. The ornamentations of the scales have been used to discriminate between congeneric *Mugil* species (Ibañez and Gallardo-Cabello 2005); likewise, the roughness of scale surfaces has been studied from the viewpoint of hydrodynamics (Sudo et al. 2002). Thus, scale shape variations possibly arise from adaptations to varying hydrodynamic conditions, and further analyses such as the present one combined with biomechanical analyses offer promise in approaching the study of form–function relationships.

Although the scales used in this study were sampled from the same anatomical region of the fish (the shoulder region), some variation exists within this region in each fish. Further, there is considerable variation over the whole body (A. Ibañez, unpublished data). The localised variations doubtless add to error by artificially inflating intragroup variability. This could be reduced by being more specific with regard to line and row of scales in future studies. The variations over the whole body might, in turn, be applied to improve fish identification rates because these larger scale variations might also be informative.

Finally, the analyses presented involved shape rather than size and shape. The centroid size of each scale was not included in the discriminant analyses. This should be factored into the analysis to test whether it improves or worsens discrimination. Considering the efficacy of discriminating stocks using scale shape as assessed by landmarks, it is somewhat surprising that the methodology has not been developed more strongly in fisheries science. This approach seems potentially useful in discriminating between closely related species that are otherwise difficult to distinguish, between fish of hybrid origin, and between conspecifics in polyploid populations (e.g., Martins et al. 1998; Ribeiro et al. 2003). The method is nondestructive, thus allowing rare and endangered species to be returned to the water. It is also quick and less costly than genetic analysis, thus allowing many individuals from a population or community to be screened quickly.

Acknowledgements

We could not have finished this study without the kind help of Manos Koutrakis, who supplied the fish scales from the Greek mullets. We also thank Markus Bastir, who re-

© 2007 NRC Canada

solved methodological and software questions for the lead author in her own language.

References

Arnold, E.L., and Thompson, J.R. 1958. Offshore spawning of the striped mullet *Mugil cephalus* in the Gulf of Mexico. Copeia, 1958: 130–132.

Bookstein, F.L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans. Pattern Anal. Machine Intel. **11**: 567–585.

Bookstein, F.L., Connor, P.D., Covell, K.D., Barr, H.M., Gleason, C.A., Sze, R.W., McBroom, J.A., and Streissguth, A.P. 2005. Preliminary evidence that prenatal alcohol damage may be visible in averaged ultrasound images of the neonatal human corpus callosum. Alcohol, **36**: 151–160.

Cadrin, S.X. 2000. Advances in morphometric identification of fishery stocks. Rev. Fish Biol. Fish. **10**: 91–112.

Caldara, F., Bargelloni, L., Ostellari, L., Penzo, E., Colombo, L., and Patarnello, T. 1996. Molecular phylogeny of grey mullets based on mitochondrial DNA sequence analysis: evidence of a differential rate of evolution at the intrafamily level. Mol. Phylogenet. Evol. **6**: 416–424.

Casselman, J.M., Collins, J.J., Crossman, E.J., Ihssen, P.E., and Spangler, G.R. 1981. Lake whitefish (*Coregonus clupeaformis*) stocks of the Ontario waters of Lake Huron. Can. J. Fish. Aquat. Sci. **38**: 1772–1789.

Cataudella, S., Civitelli, M.V., and Capanna, E. 1974. Chromosome complement of the Mediterranean mullets (Pisces, Perciformes). Caryologia, **27**: 93.

Chen, X.M., Milne, N., and O'Higgins, P. 2005. Morphological variation of the thoracolumbar vertebrae in Macropodidae and its functional relevance. J. Morphol. **266**: 167–181.

Crews, S.C., and Hedin, M. 2006. Studies of morphological and molecular phylogenetic divergence in spiders (Araneae: Homalonychus) from the American southwest, including divergence along the Baja California Peninsula. Mol. Phylogenet. Evol. **38**: 470–487.

Crosetti, D., Nelson, W.S., and Avise, J.C. 1994. Pronounced structure of mitochondrial DNA among populations of the circumglobally distributed grey mullet (*Mugil cephalus*). J. Fish Biol. **44**: 47–58.

Ditty, J.G., and Shaw, R.F. 1996. Spatial and temporal distribution of larval striped mullet (*Mugil cephalus*) and white mullet (*M. curema*, family: Mugilidae) in the northern Gulf of Mexico, with notes on mountain mullet, *Agonostomus monticola*. Bull. Mar. Sci. **59**: 271–288.

Dryden, I.L., and Mardia, K.V. 1993. Multivariate shape analysis. Sankya Ser. A, **55**: 460–480.

Dryden, I.L., and Mardia, K.V. 1998. Statistical shape analysis. John Wiley and Sons, London, UK.

Félix-Uraga, R., Quiñones-Velázquez, C., Hill, K.T., Gómez-Muñoz, V.M., Melo-Barrera, F.N., and García-Franco, W. 2005. Pacific sardine (*Sardinops sagax*) stock discrimination off the west coast of Baja California and southern California using otolith morphometry. CCOFI Rep. **46**: 113–121.

Funicelli, N.A., Meineke, D.A., Bryant, H.E., Dewey, M.R., Ludwig, G.M., and Mengel, L.S. 1989. Movements of striped mullet, *Mugil cephalus*, tagged in Everglades National Park, Florida. Bull. Mar. Sci. **44**: 171–178.

Guinea, J., and Fernández, F. 1992. Morphological and biometrical study of the gill rakers in four species of mullets. J. Fish Biol. **41**: 381–397.

Hutchinson, W.F., Carvalho, G.R., and Rogers, S.I. 2001. Marked genetic structuring in localised spawning populations of cod *Gadus morhua* in the North Sea and adjoining waters, as revealed by microsatellites. Mar. Ecol. Prog. Ser. **223**: 251–260.

Ibañez, A.L., and Gallardo-Cabello, M. 2004. Reproduction of *Mugil cephalus* and *M. curema* (Pisces: Mugilidae) from a coastal lagoon to the northwest of the Gulf of Mexico. Bull. Mar. Sci. **75**: 37–49.

Ibañez, A.L., and Gallardo-Cabello, M. 2005. Identification of two Mugilidae species, *Mugil cephalus* and *M. curema* (Pisces: Mugilidae), using the ctenii of their scales. Bull. Mar. Sci. **77**: 305–308.

Ibañez, A.L., and Gutiérrez-Benítez, O. 2004. Climatic variables and spawning migrations of the striped mullet and white mullet in the northwestern area of the Gulf of Mexico. J. Fish Biol. **65**: 822–831. doi:10.1111/j.1095–8649.

Idyll, D.P., and Sutton, J.W. 1951. Results of the first years tagging of mullet *Mugil cephalus* L. on the west coast of Florida. Trans. Am. Fish. Soc. **81**: 69–77.

Ihssen, P.E., Brooke, H.E., Casselman, J.M., McGlade, J.M., Payne, N.R., and Utter, F.M. 1981. Stock identification: materials and methods. Can. J. Fish. Aquat. Sci. **38**: 1838–1855.

Jarvis, R.S., Klodowski, H.F., and Sheldon, S.P. 1978. New method of quantifying scale shape and an application to stock identification in walleye (*Stizostedion vitreum vitreum*). Trans. Am. Fish. Soc. **107**: 528–534.

Jerry, D.R., and Cairns, S.C. 1998. Morphological variation in the catadromous Australian bass, from seven geographically distinct riverine drainages. J. Fish Biol. **52**: 829–843.

Kent, J.T. 1994. The complex Bingham distribution and shape analysis. J. R. Statist. Soc. B, **56**: 285–299.

Kinsey, S.T., Orsoy, T., Bert, T.M., and Mahmoudi, B. 1994. Population structure of the Spanish sardine *Sardinella aurita*: natural morphological variation in a genetically homogeneous population. Mar. Biol. **118**: 309–317.

Koutrakis, E.T. 2004. Temporal occurrence and size distribution of grey mullet juveniles (Pisces, Mugilidae) in the estuarine systems of the Strymonikos Gulf (Greece). J. Appl. Ichthyol. **20**: 76–78.

Koutrakis, E.T., Sinis, A.I., and Economidis, P.S. 1994. Seasonal occurrence, abundance and size distribution of gray mullet fry (Pisces, Mugilidae) in the Porto-Lagos Lagoon and Lake Vistonis (Aegean Sea, Greece). Bamidgeh, **46**: 182–196.

Mahmoudi, B. 1991. Population assessment of black mullet in the eastern Gulf of Mexico. Final report of Cooperative Agreement (MARFIN) NA90-WC-H-W003, Florida Department of Environmental Protection, St. Petersburg, Fla.

Marcus, L.F., Corti, M., Loy, A., Naylor, G.J.P., and Slice, D.E. (*Editors*). 1996. Advances in morphometrics. NATO ASI, Ser. A, Life Sciences, Vol. 284. Plenum Press, New York.

Mariani, S., Hutchinson, W.F., Hatfield, E.M.C., Ruzzante, D.E., Simmonds, E.J., Dahlgren, T.G., Andre, C., Brigham, J., Torstensen E., and Carvalho, G.R. 2005. North Sea herring population structure revealed by microsatellite analysis. Mar. Ecol. Prog. Ser. **303**: 245–257.

Martins, M.J., Collares-Pereira, M.J., Cowx, I.G., and Coelho, M.M. 1998. Diploids versus triploids of *Rutilus alburnoides*: morphological distinction and habitat preferences. J. Fish Biol. **52**: 817–828.

Monteiro, L.R., Di Beneditto, A.P.M., Guillermo, L.H., and Rivera, L.A. 2005. Allometric changes and shape differentiation of sagitta otoliths in sciaenid fishes. Fish. Res. **74**: 288–299.

Murta, A.G. 2000. Morphological variation of horse mackerel (*Trachurus trachurus*) in the Iberian and North African Atlantic: implications for stock identification. ICES J. Mar. Sci. **57**: 1240–1248.

© 2007 NRC Canada

O'Higgins, P. 2000. Advances in approaches to the study of morphological variation in the hominid fossil record: biology, landmarks and geometry. J. Anat. **197**: 103–120.

O'Higgins, P., Chadfield, P., and Jones, N. 2001. Facial growth and the ontogeny of morphological variation within and between the primates *Cebus apella* and *Cercocebus torquatus*. J. Zool. **254**: 337–357.

Papasotiropoulos, V., Klossa-Kilia, E., Kilias, G., and Alahiotis, S. 2001. Genetic divergence and phylogenetic relationships in grey mullets (Teleostei: Mugilidae) using allozyme data. Biochem. Genet. **39**: 155–168.

Papasotiropoulos, V., Klossa-Kilia, E., Kilias, G., and Alahiotis, S. 2002. Genetic divergence and phylogenetic relationships in grey mullets (Teleostei: Mugilidae) based on PCR–RFLP analysis of mtDNA segments. Biochem. Genet. **40**: 71–86.

Poulet, N., Reyjol, Y., Collier, H., and Lek, S. 2005. Does fish scale morphology allow the identification of populations at a local scale? A case study for rostrum dace *Leuciscus leuciscus burdigalensis* in River Viaur (SW France). Aquat. Sci. **67**: 122–127.

Recasens, L., Lombarte, A., and Sanchez, P. 2006. Teleostean fish assemblages in an artificial reef and a natural rocky area in Catalonia (northwestern Mediterranean): an ecomorphological approach. Bull. Mar. Sci. **78**: 71–82.

Ribeiro, F., Cowx, I.G., Tiago, P., Filipe, A.F., Moreira da Costa, L., and Collares-Pereira, M.J. 2003. Growth and reproductive traits of diploid and triploid forms of the *Squalius alburnoides* cyprinid complex in a tributary of the Guadiana River, Portugal. Arch. Hydrobiol. **156**: 471–484.

Richards, R.A., and Esteves, C. 1997. Use of scale morphology for discriminating wild stocks of Atlantic striped bass. Trans. Am. Fish. Soc. **126**: 919–925.

Rocha-Olivares, A., Garber, N.M., and Stuck, K.C. 2000. High genetic diversity, large inter-oceanic divergence and historical demography of the striped mullet. J. Fish Biol. **57**: 1134–1149.

Rossi, A.R., Capula, M., Crosetti, D., Sola, L., and Campton, D.E. 1998. Allozyme variation in global populations of *Mugil cephalus* (Pisces: Mugilidae). Mar. Biol. **131**: 203–212.

Rossi, R. 1986. Occurrence, abundance and growth of fish fry in Scardovari Bay, a nursery ground of the Po River delta (Italy). Arch. Oceanogr. Limnol. **20**: 259–280.

Shepherd, G. 1991. Meristic and morphometric variation in black sea bass north of Cape Hatteras, North Carolina. N. Am. J. Fish. Manag. **11**: 139–148.

Sudo, S., Tsuyuki, K., Ito, Y., and Ikohagi T. 2002. A study on the surface shape of fish scales. JSME Int. J. Ser. C, **45**: 1100–1105.

Swain, D.P., and Foote, C.J. 1999. Stocks and chameleons: the use of phenotypic variation in stock identification. Fish. Res. **43**: 113–128.

Thompson, B.A., Render, J.H., Allen, R.L., and Nieland, D.L. 1991. Fisheries independent characterization of population dynamics and life history of striped mullet in Louisiana. Coastal Fisheries Institute, Final Report to the US Department of Commerce – National Marine Fisheries Service, Louisiana State University CFI-90-01. pp. 1–92.

Thomson, J.M. 1981. The taxonomy of grey mullets. *In* Aquaculture of grey mullets. *Edited by* O.H. Oren. Cambridge University Press, Cambridge. pp. 1–16.

Turan, C., Caliskan, M., and Kucuktas, H. 2005. Phylogenetic relationships of nine mullet species (Mugilidae) in the Mediterranean Sea. Hydrobiologia, **532**: 45–51.

Tzanatos, E., Dimitriou, E., Katselis, G., Georgiadis, M., and Koutsikopoulos, C. 2005. Composition, temporal dynamics and regional characteristics of small-scale fisheries in Greece. Fish. Res. **73**: 147–158.

Walker, J.A. 1996. Principal components of body shape variation within an endemic radiation of threespine stickleback. *In* Advances in morphometrics. *Edited by* L.F. Marcus, M. Corti, A. Loy, G.J.P. Naylor, and D.E. Slice. NATO ASI Ser. A, Life Sciences, Vol. 284, Plenum Press, New York. pp. 321–334.

Watkinson, D.A., and Gillis, D.M. 2005. Stock identification of Lake Winnipeg walleye based on Fourier and wavelet description of scale outline signals. Fish. Res. **72**: 193–203.

West-Eberhard, M.J. 1989. Phenotypic plasticity and the origins of diversity. Annu. Rev. Ecol. Syst. **20**: 249–278.

© 2007 NRC Canada

Integrating multiple sources of data on migratory timing and catchability to estimate escapement for steelhead trout (*Oncorhynchus mykiss*)

Josh Korman, Caroline C. Melville, and Paul S. Higgins

Abstract: We evaluated the influence of biological, physical, and year effects on catchability, survey life, and departure timing for a winter-run steelhead (*Oncorhynchus mykiss*) population and compared the performance of alternate escapement models. Date of entry and gender explained 65% of the variability in survey life, and there was no evidence for differences in survey life among years. The median date of departure for male spawners occurred 2 weeks later relative to females, and a gender-based departure model was strongly supported. Departure timing was significantly different among years ($p < 0.05$), but the maximum difference in median departure dates was only 11 days. The ratio of horizontal visibility to discharge explained 50% of the variation in catchability, and there was weak support for a model that accounted for effects associated with courtship and spawning behaviour. There was strong support for an escapement model that assumed survey life and catchability relationships were common among years. Joint use of departure timing and survey life data reduced uncertainty in escapement estimates by an average of 40%. The major advantage of our escapement model is that it increases the precision of estimates while avoiding the use of overly simplistic assumptions about run timing and catchability.

Résumé : Nous évaluons l'influence des effets des facteurs biologiques, des facteurs physiques et de l'année sur la capturabilité, la durée de la présence dans la zone d'inventaire et le calendrier de départ chez une population de truites arc-en-ciel anadromes (*Oncorhynchus mykiss*) à montaison hivernale et comparons les performances de divers modèles de rechange de l'échappement. La date d'entrée et le sexe expliquent 65 % de la variabilité de la durée de présence dans la zone d'inventaire et il n'y a aucune indication de différence dans cette durée entre les années. La date médiane de départ des reproducteurs mâles se situe 2 semaines plus tard par rapport à celle des femelles; cela appuie fortement le modèle de départ basé sur le sexe. Le calendrier des départs diffère significativement d'une année à l'autre ($p < 0,05$), mais la différence maximale entre les dates médianes de départ est de seulement 11 jours. Le rapport de la visibilité horizontale sur le débit explique 50 % de la variation de la capturabilité; il y a un appui faible pour un modèle qui tienne compte des effets associés aux comportements de cour et de fraye. Il y a cependant un fort appui pour un modèle d'échappement qui assume que les relations entre la durée de présence dans la zone d'inventaire et la capturabilité sont les mêmes d'une année à l'autre. L'utilisation conjointe du calendrier de départ et des données de présence dans la zone d'inventaire réduit l'incertitude dans les estimations de l'échappement de 40 % en moyenne. L'avantage majeur de notre modèle d'échappement est qu'il améliore la précision des estimations, tout en évitant de faire des présuppositions trop simplistes sur le calendrier de départ et la capturabilité.

[Traduit par la Rédaction]

Introduction

Annual escapement estimates for adult salmon and steelhead (*Oncorhynchus mykiss*) populations are routinely used to assess stock status, habitat impacts, and restoration efforts. Although there is widespread concern about the status of wild steelhead populations (Slaney et al. 1996), there is generally a paucity of reliable escapement estimates to evaluate the extent of the problem and the efficacy of recovery efforts. Steelhead typically spawn in higher-order streams that are too large for fish-counting weirs, and their protracted return timing leads to high operating costs in systems where weirs can be used. Implementation of repeat mark–recapture experiments is the only feasible method to estimate escapement in most circumstances. However, as steelhead populations are generally much less numerous than salmon populations, estimates will usually be very imprecise because of low numbers of recapture events spread out over a protracted run. Improvements in steelhead enumeration will require the use of tagging technologies that provide as much

Received 22 November 2006. Accepted 19 April 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 8 August 2007. J19674

J. Korman.[1] Ecometric Research Inc. and Department of Zoology, University of British Columbia, 3560 W 22nd Avenue, Vancouver, BC V6S 1J3, Canada.
C.C. Melville. Instream Fisheries Research Inc. 223, 2906 West Broadway, Vancouver, BC V6K 2G8, Canada.
P.S. Higgins. BC Hydro and Power Authority, 6911 Southpoint Drive, Burnaby, BC V3N 4X8, Canada.

[1]Corresponding author (e-mail: jkorman@ecometric.com).

 doi:10.1139/F07-083 © 2007 NRC Canada

information as possible and analytical methods that make full use of these data.

Escapement is often estimated by obtaining repeat counts of the number of fish present over the spawning migration. The proportion of fish observed (catchability) and the proportion of the total run that is present must both be estimated on each survey to determine the total escapement. Catchability, often referred to as observer efficiency in the context of stream surveys, can be determined through mark–recapture experiments. The total number of fish present on a survey is simply the difference between the cumulative arrivals and departures by the survey date. The departure schedule will be determined based on the arrival schedule and the length of time that spawners remain in the survey area (survey life). The proportion of the run that is present on any survey date can therefore be estimated from data on at least two of three run-timing components: arrival timing, survey life, and departure timing. A better understanding of the physical and biological processes that influence run timing and catchability and the extent of interannual variability will lead to more informed survey designs and improved escapement estimates.

In many rivers, catchability of visual-based surveys will be temporally heterogeneous, especially for winter-run steelhead where migration timing occurs over a 5-month period when there are very large differences in flow and water clarity (Korman et al. 2002). Catchability may be spatially heterogeneous within the survey area because of differences in river morphology and tributary inputs of fine sediment. If within-river migratory behaviour changes the proportion of fish in these different environments across surveys, the overall extent of temporal variation in catchability will increase. Fish entering a stream well before spawning tend to have longer survey lives compared with fish that arrive later (e.g., Su et al. 2001; Korman et al. 2002; Quinn 2005). Predictions of survey life based on date of entry are generally imprecise, in part because important physical and biological covariates are not taken into account. The effect of interannual variation in survey life on escapement estimates has not been evaluated but has been recognized as a potentially important source of error in escapement estimates (Adkison and Su 2001).

Analytical approaches for estimating escapement from repeat count data have advanced considerably from the original area-under-the-curve (AUC) methodology (e.g., English et al. 1992). Hilborn et al. (1999) used a maximum-likelihood approach to estimate escapement and arrival-timing parameters by assuming that survey life was constant and that, on average, all fish present in the survey area were counted. Korman et al. (2002) estimated escapement from repeat mark–recapture experiments in conjunction with more flexible arrival timing and survey life models. Escapement estimates will be uncertain if there are no counts after the peak (Hilborn et al. 1999; Adkison and Su 2001) or if peak and after-peak surveys occur during periods of low catchability (Korman et al. 2002). In these situations, the possibility of a large number of fish entering at the peak or late in the run cannot be discounted in the estimation process because there is little information about arrival timing in the repeat-count data. This problem becomes more severe with increases in the flexibility of arrival timing and survey life model structures. Su et al. (2001) used a hierarchical Bayesian model to "borrow" information on arrival timing from years when it was better defined to improve estimates of escapement and timing in years when few or no counts were made after the peak of escapement. The approach has utility for escapement programs with many years of data and timing that is well defined in at least some years, but it is not very useful in programs with less data and (or) escapement timing that is never well defined because of deterioration in catchability at the peak or end of the run.

In this paper, we develop an improved method for estimating escapement from repeat-count data that addresses many of the limitations and uncertainties of past efforts. We analyze a multiyear radiotelemetry and snorkel count data set for a steelhead population that spawns in the Cheakamus River, British Columbia, to evaluate the extent of interannual variability and the importance of physical and behavioural factors on catchability, survey life, and departure timing. The escapement estimation model presented in Korman et al. (2002) is considerably revised to better account for uncertainty in survey life and catchability estimates and to incorporate information on departure timing. Survey life – date of entry parameters are jointly estimated along with all other model parameters by including an independent term for survey life data in the overall likelihood. This differs from the approaches developed by Hilborn et al. (1999), Su et al. (2001), and Korman et al. (2002) in which uncertainty in survey life was ignored. Su et al. (2001) used a Bayesian hierarchical model to improve precision of arrival timing and escapement estimates. Here we adopt a different approach by including data on departure schedules obtained by radiotelemetry as another independent likelihood term. Korman et al. (2002) estimated survey-specific catchabilities from mark–recapture data or a static physical model predicting catchability based on river conditions in years when no tagging data were available. The escapement model presented here includes data on catchability from both sources in the overall likelihood and accounts for uncertainty in the physical model. The model is applied to the historical Cheakamus River swim count data set, which consists of years with and without telemetry data, to estimate a 10-year escapement time series. We examine the sensitivity of the time series to individual likelihood components and alternate assumptions about the exchangeability of survey life and catchability data across years.

Materials and methods

Swim and telemetry data

Returning steelhead spawners to the Cheakamus River, British Columbia, have been enumerated by snorkel survey since 1996 (1996–2006, excluding 1998). Three divers distributed equidistant along the river cross section float the upper 14.5 km of stream length accessible to anadromous fish and record the number steelhead observed in upper (river km 4.5–14.5) and lower (river km 0–4.5) sections of the survey area (see fig. 1 of Korman et al. 2002). On average, nine surveys have been conducted per year, typically between late February and mid- to late May. Total counts per survey were low and averaged 36 fish (1 standard deviation (SD) = 21). A radiotelemetry program was conducted in 5 of

© 2007 NRC Canada

the 10 survey years (2000–2005, excluding 2002). The program provided information on observer efficiency in all 5 years; however, data on survey life and departure timing were only obtained in the latter 4 years. Fish were captured by angling and fitted with both radio and external tags, and the number and location of marked fish observed by divers during their normal surveys were recorded. The locations of all radio-tagged steelhead were determined using data from fixed stations located at the downstream ends of the upper and lower survey sections and by mobile tracking during the swim surveys. A total of 143 tags were deployed over 5 years. The majority of fish that were tagged were captured within the survey area. Survey life could only be determined for the 33 fish that were caught downstream and subsequently entered the survey area. The date of downstream departure from the survey area was determined for 104 fish that included the 33 fish for which date of entry was also determined. See Korman et al. (2002) for further details on the study area, swim survey, and telemetry methodologies.

Data analysis

We combined relevant telemetry data from 2000–2005 and used a maximum-likelihood approach (Hilborn and Mangel 1997) to examine (*i*) alternate structural forms of the survey life – date of entry relationship and the effects of covariates such as gender and year; (*ii*) alternate structural forms for models predicting catchability based on horizontal visibility, discharge, survey date, water temperature, and the spatial distribution of tagged fish; and (*iii*) the effects of year and gender on departure timing. The analysis of departure timing consisted of a set of nested models; thus a likelihood ratio test was used to determine if the additional variance explained by more complex models outweighed the predictive cost associated with the additional parameters. Twice the difference between the log-likelihoods of two models is χ^2 distributed, with degrees of freedom equal to the difference in the number of parameters between models (Hilborn and Mangel 1997).

We used the second-order Akaike information criterion (AIC_c) to compare the out-of-sample predicted power of alternate survey life and catchability models (Burnham and Anderson 2002). AIC_c scores were computed as the sum of twice the deviance at the maximum-likelihood estimate and the number of parameters adjusted by a correction factor that accounts for small sample size. The AIC_c criterion describes the trade-off between model fit and complexity. Models with the lowest scores are those with the fewest number of parameters that still provide an adequate representation of the data. Models with too few parameters will provide poor fits to the data because they fail to recognize important structure. Models with too many parameters may fit the data better but will have variances that are needlessly large and may identify spurious structures. The difference between the lowest AIC_c score across a set of candidate models and the score for each model, referred to as ΔAIC, measures the extent of support for each model given the data. We use the ΔAIC guidelines for model support provided by Burnham and Anderson (2002): $\Delta AIC < 2$ (substantial or strong support); $4 < \Delta AIC < 7$ (weak support); $\Delta AIC > 10$ (essentially no support).

Error structures used in model evaluations were consistent with those from the escapement estimation model. We assumed that errors for relationships predicting survey life and catchability followed normal and Poisson distributions, respectively. The Poisson error model accounts for differences in the amount of information in each catchability estimate. Estimates of catchability when few tags are present or when the probability of seeing a tag is very low are much more uncertain than estimates obtained under the opposite conditions. An inverse logit transformation of predicted catchabilities was used to constrain values between 0 and 1 for linear models. To examine variation in departure timing by gender and year, we fit cumulative beta distributions to the data assuming multinomial error. Most likely parameter estimates were computed for all models using a nonlinear iterative search procedure that maximized the sum of log-transformed normal, Poisson, or multinomial probabilities across all observations.

Escapement model structure and estimation

The escapement model consists of three main elements. A simple process model predicts the number of fish present on each day of the run and the departure schedule based on the total escapement and parametric relationships simulating arrival timing and survey life. An observation model simulates the number of marked and unmarked fish observed on each survey based on the number of tags known to be in the survey area and predictions of the number of unmarked fish that are present and catchability. Process and observation model parameters are estimated by maximizing the value of a likelihood function that integrates data on the number of marked and unmarked fish observed on each survey and survey life and departure schedules measured by radiotelemetry.

Process model

The proportion of the total escapement entering the survey area on day i (PA_i) of the simulated run is predicted by a beta distribution

$$(1) \qquad PA_i = \phi_i^{\tau-1}(1-\phi_i)^{\beta-1}$$

where τ and β are parameters of the beta distribution and ϕ_i represents the proportional day of the run for day i, ranging from 0 to 1 on the assumed first (January 1, $i = 1$) and last (June 30, $i = 181$) day, respectively. τ is the precision of run timing, with smaller values representing a low and constant rate of arrival over the duration of the run and larger values representing a shorter and more concentrated arrival timing. The beta distribution is reparameterized so that β is calculated based on estimates of the day when the peak arrival rate occurs (μ) and the precision of arrival timing, using the transformation $\beta = (\tau - 1)(\mu / T)^{-1} + 2 - \tau$ (Gelman et al. 2004, their appendix A).

Survey life, the number of days that a fish spends in the survey area, is predicted using a negative logistic relationship:

$$(2) \qquad SL_i = \lambda_m\left(1 - \frac{i^{\lambda_s}}{\lambda_h^{\lambda_s} + i^{\lambda_s}}\right)$$

© 2007 NRC Canada

where SL_i is the mean survey life for a fish entering on day i, λ_m is the maximum mean survey life, λ_b is the day at which mean survey life is half the maximum, and λ_s is the slope of the relationship. The mean departure day for a fish arriving on day i is simply $d_i = i + SL_i$.

The proportion of fish that arrive on day i and depart on day j ($PAD_{i,j}$) is predicted from a normal distribution with mean d_i and standard deviation σ_d.

$$(3) \quad PAD_{i,j} \sim \text{Normal}(j, d_i, \sigma_d)$$

PAD values are standardized so that proportions across all departure days for each arrival day sum to 1, that is, all fish must exit the survey area by the assumed last day of the run. As a fish cannot depart before it arrives, $PAD_{i,j} = 0$ for $j < i$. The proportion of fish departing on each day (PD_j) is computed from

$$(4) \quad PD_j = \sum_i PA_i PAD_{i,j}$$

Note that departure timing depends on both arrival timing and the survey life relationship that defines PAD. Finally, the number of unmarked fish present in the survey area on each day (U_i) is the product of the total escapement (E) less the number of fish marked per year (R) and the difference between the cumulative arrivals and departures of unmarked fish on that day:

$$(5) \quad U_i = (E - R)\left[\sum_1^i PA - \sum_1^i PD\right]$$

The difference between the cumulative values of PA and PD on any date represents the proportion of the total run that is present.

Observation model and estimating model parameters

Escapement, arrival timing, and survey life parameters and those defining the relationship between catchability (q) and the ratio of horizontal visibility to discharge (HV/Q) are jointly estimated by maximum-likelihood or Bayesian methods. Independent likelihood terms are developed for counts of marked and unmarked fish, survey life, departure schedule, and q–HV/Q data. The log-likelihoods are added together to give an overall likelihood function.

The likelihoods of the number of marked (L_r) and unmarked (L_u) fish observed are assumed to follow a Poisson distribution:

$$(6) \quad r_i \sim \text{Poisson}(q_i R_i)$$

$$(7) \quad u_i \sim \text{Poisson}(q_i U_i)$$

where R_i is the total number of marked fish present for a survey conducted on day i (determined from telemetry data), r_i is the number of marked fish that are observed by swimmers, U_i is the predicted number of unmarked fish that are present (from eq. 5), u_i is the number of unmarked fish observed, and q_i is the catchability coefficient. The terms L_r and L_u, as for all that follow, represent the sum of log-transformed probabilities across all observations. Note that catchability is a nuisance parameter that does not need to be directly estimated. It is evaluated at its conditional maximum-likelihood estimate for each survey using (Korman et al. 2002):

$$(8) \quad q_i = \frac{r_i + u_i}{R_i + U_i}$$

That is, catchability is simply the ratio of the total number of fish observed to the total number present. Values of U_i are not independent across surveys because they are linked through the model structure, thus the number of unmarked fish observed contributes to the estimate of catchability.

Korman et al. (2002) found that the ratio of horizontal visibility to discharge on each survey was a reasonable predictor of catchability, calculated as the ratio of tags observed to tags present. Physically based catchability predictions are required to estimate the number of fish present in years when there is no tagging. In this analysis, we recognize that catchability predictions can also be used in years with tagging information to increase the precision of estimates of the numbers present. Precision of a purely tag-based estimate of catchability will be very poor when the total number of tags present or the true catchability is very low. In this situation, estimates of catchability from the physically based model, which incorporates information on catchability from multiple surveys within and across years, will make an important contribution to the estimate of the numbers present. The following logistic model is used to predict catchability based on river conditions in the escapement estimation procedure:

$$(9) \quad qP_i = \frac{\frac{HV^{\rho_s}}{Q}}{\rho_b^{\rho_s} + \frac{HV^{\rho_s}}{Q}}$$

where qP_i is the physically based prediction of catchability, HV/Q is the ratio of horizontal visibility to discharge, ρ_b is the HV/Q ratio at which observer efficiency is 0.5, and ρ_s is the slope of the relationship. Two additional likelihoods for the observed number of marked (L_{rp}) and unmarked (L_{up}) fish can now be computed by replacing the conditional catchabilities in eqs 6 and 7 (q_i) with catchabilities predicted by the physical model (eq. 9).

Parameters of the qP–HV/Q relationship are jointly estimated with other model parameters using data from all surveys when tags were present in all years when telemetry was conducted ($n = 52$ across 5 years). The likelihood of the observed number of marked fish (L_p) on each of these surveys is assumed to follow a Poisson distribution:

$$(10) \quad r_i \sim \text{Poisson}(qP_i R_i)$$

Note that L_{rp} is the sum of likelihoods across surveys in only the year in which escapement is being estimated. L_p is the sum of likelihoods across all surveys when tags were present over all years in which telemetry was conducted, excluding observations used in calculating L_{rp} (to avoid double-counting).

The likelihood of the survey life data (L_s) was computed assuming normally distributed error

$$(11) \quad \text{slobs}_i \sim \text{Normal}(i, SL_i, \sigma_d)$$

where slobs_i is the observed survey life for a fish that entered the survey area on day i, and SL_i is the predicted survey life for the same fish based on its observed date of entry

© 2007 NRC Canada

(eq. 2). Note that σ_{sl} is a nuisance parameter that is calculated at its conditional maximum-likelihood value of $\sqrt{\frac{\sum(\text{slobs}_i - \text{SL}_i)^2}{n-1}}$, where n ($n = 33$) is the number of survey life observations (Walters and Ludwig 1994).

The likelihood of the observed number of fish exiting the lower survey area in a downstream direction (departure schedule – L_d) was computed assuming multinomial error,

$$(12) \quad \text{nexit}_i \sim \text{Multinom}(\text{texit}, \text{PD}_i)$$

where texit is the total number of radio-tagged fish for which an exit date could be determined ($n = 104$), nexit_i is the number of radio-tagged fish that departed on day i, and PD_i is the predicted departure proportions for that day (from eq. 4).

The overall log-likelihood of all the data given a set of model parameters $\theta = E, \mu, \tau, \lambda_m, \lambda_h, \lambda_s, \rho_h, \rho_s$ was determined by summing all component log-likelihoods:

$$(13) \quad L_T(\text{data}|\theta) = \frac{L_r + L_{gr}}{2} + \frac{L_u + L_{gu}}{2} + L_p + L_s + L_d$$

When estimating parameters for a particular year, note that the first four terms of the total likelihood are evaluated based only on data collected in that year, whereas the latter three depend on data collected over all years when telemetry was conducted. The denominator of 2 in eq. 13 accounts for the fact that observations of marked and unmarked fish are double-counted in the overall likelihood because they are evaluated using both conditional MLE values (q from eq. 8) and physically based predictions of catchability (qP from eq. 9). The first term of eq. 13 does not contribute to the total likelihood in years with no tagging or for surveys in which no tags are present in years when tagging is conducted. In years when tagging data are available, highest likelihoods of the objective function will occur when escapement, arrival, and survey life parameters maximize the survey-specific likelihoods but also predict a survey life relationship and departure schedule that is consistent with observations collected across all telemetry years and a set of catchabilities consistent with predictions from the physically based model, which again is dependent on data from all telemetry years.

The overall likelihood model assumes that data relating to catchability, survey life, and departure timing are exchangeable among years and that year-specific tagging data are available. To evaluate these assumptions, various terms from eq. 13 (likelihood model 1 in Table 1) were removed prior to fitting the model. Like the situation for the Cheakamus program, radiotelemetry data may only be available for a subset of years over which counts have been conducted. In the majority of years, parameters need to be estimated in the absence of tagging information (likelihood model 2 in Table 1, remove L_r and L_{gr}). Likelihoods associated with the physical model predicting catchability were removed (remove L_{gr}, L_{gu}, L_{gs}) in years when fish were tagged to address concerns about the exchangeability of catchabilities among years, perhaps as a result of changes in river morphology or observers (likelihood model 3). It may be difficult or impossible to measure survey life in some cases, if, for example, one cannot catch fish downstream of the survey area (likelihood model 4, remove L_s). The exchangeability of departure-timing data among years would be a concern if there is strong interannual variation in the survey life – date of entry relationship or arrival timing (likelihood model 5, remove L_d). The overall likelihood can be reduced to the simpler versions used in Korman et al. (2002) by excluding components for physical predictions of catchability, survey life, and departure timing (likelihood 6 and 7). Note also that Korman et al. (2002) did not use physical predictions of catchability to evaluate the likelihood of the number of tagged (L_{gr}) and untagged (L_{gu}) fish observed in years when tags were present (likelihood 6).

Two alternate hierarchies were used to estimate a 10-year time series of escapement for the Cheakamus River steelhead population ($E_{1996} - E_{2006}$). In the first case, all model parameters were estimated independently for each year from 1996–2006. Eight parameters were separately estimated for each of the 10 years for which count data were available, for a total of 80 parameters ($k = 80$). In the second case, parameters for survey life – date of entry and qP–HV/Q relationships were assumed to be common to all years, whereas escapement and arrival-timing parameters were not. In this case, 30 year-specific parameters (E, μ, τ) and five parameters common to all years were estimated ($k = 35$ for all 10 survey years). To summarize, two alternate estimation structures were explored:

$$(14) \quad \text{Year-independent}\,(k=80): \quad \theta_{year} = E_{-year}, \mu_{-year}, \tau_{-year}, \lambda_{m-year}, \lambda_{h-year}, \lambda_{s-year}, \rho_{h-year}, \rho_{s-year}$$

$$(15) \quad \text{Common year}\,(k=35): \quad \theta_{year-c} = E_{-year}, \mu_{-year}, \tau_{-year}, \lambda_m, \lambda_h, \lambda_s, \rho_h, \rho_s$$

Most likely parameter estimates (MLEs) were computed using the AD model builder software (Otter Research 2004). Asymptotic estimates of the standard deviation for the MLEs were computed using the delta method. Joint posterior distributions of model parameters were estimated using Markov chain Monte Carlo simulation (MCMC). We used uniform priors for all model parameters and selected bounds that were wide enough to have no effect on the estimation of the joint posteriors. Joint posterior distributions were based on 100 000 simulations for each set of year-specific estimates for the year-independent model, or 100 000 simulations for all years for the common-year model. Two independent Markov chains were simulated to estimate joint posterior distributions. The chains had highly divergent initial parameter values to represent the possibilities of high and low escapements and early and late run timings and were initiated at different random number seeds. Every 25th simulation from each chain was saved to a text file and imported to the R statistical package (http://www.r-project.org). A series of diagnostics implemented in the R CODA package (Convergence Diagnostics and Output Analysis, http://www-fis.iarc.fr/coda/; see appendix B of Su et al. (2001) for a description of CODA) were then applied to the sample of 4000 parameter values from each chain to evaluate within- and across-chain convergence. The diagnostics included Gelman and Rubin's potential scale reduction factor, Geweke's test for equality of means, and Heidelberger and Welch's and

© 2007 NRC Canada

Table 1. Likelihood formulations to account for alternate assumptions about the exchangeability of data among years and differences in availability of data due to survey design (see eq. 13).

Likelihood model no.	Structural assumptions or data limitations	Components of objective function
1	Complete (all data available and exchangeable)	$L_T = \frac{L_u + L_t}{2} + \frac{L_{pu} + L_{pt}}{2} + L_p + L_v + L_d$
2	No tags in year of survey	$L_T = \frac{L_u + L_{pu}}{2} + L_p + L_v + L_d$
3	No qP–HV/Q relationship (must have tags in year of survey)	$L_T = L_u + L_t + L_v + L_d$
4	No survey life data	$L_T = \frac{L_u + L_t}{2} + \frac{L_{pu} + L_{pt}}{2} + L_p + L_d$
5	No departure timing data	$L_T = \frac{L_u + L_t}{2} + \frac{L_{pu} + L_{pt}}{2} + L_p + L_v$
6	Korman et al. 2002 (tags in year of survey)	$L_T = L_u + L_t$
7	Korman et al. 2002 (no tags in year of survey	$L_T = L_{pu} + L_{pt}$

Raftery and Lewis's run length control diagnostics (see Gelman et al. 2004, section 11.6 or appendix B of Su et al. 2001). The final joint posterior distributions were estimated by combining the last 2000 samples from each chain, which occurred well after convergence was established in all cases.

We used the AIC_c and deviance information criterion (DIC) to compare the out-of-sample predicted power of year-independent and common-year escapement models. DIC is a Bayesian analogue of AIC because it accounts for uncertainty in parameter estimates and model fit (Gelman et al. 2004). DIC scores were computed as the sum of the average deviance across posterior samples (D_{avg}) and the effective number of parameters (pD), with the latter statistic computed as the difference between D_{avg} and the deviance at the posterior means. pD characterizes model complexity and approximates the number of "unconstrained" parameters in the model, where a parameter counts as 1 if it is estimated with no constraints or prior information, 0 if it is fully constrained or all the information about the parameter comes from the prior distribution, or an intermediate value if both the data and prior distributions are informative (Gelman et al. 2004).

Results

Assessment of alternate structural forms for process model components

We compared a series of models predicting survey life, departure timing, and catchability to determine the most appropriate structures for the escapement model. Gender and date of entry were important determinants of survey life. There was essentially no support for survey life models that did not include date of entry as an explanatory variable (Table 2, models 1 and 2). There was strong support for the linear model predicting survey life based on entry day (model 3), and this model explained 50% of the variation in survey life. Note that linear models will likely overpredict survey life for early-entry fish and underestimate survey life for late-entry fish. We therefore evaluated the fit of an arctangent-transformed linear model with a maximum range in survey life equivalent to the assumed duration of the total run (model 4). There was very little support for this model. There was strong support for linear models with gender-specific slopes (model 5) or constants (model 6). The latter model had the greatest support of all candidate models and explained 65% of the variation in survey life (Fig. 1). There was little to no support for models that included a year effect (models 8–10). The model with gender-specific slopes and a year-specific constant (model 10) had near strong support, but note that its ΔAIC_c value was over 2 points higher than the equivalent model without a year effect (model 5). Thus adding a year-dependent slope did not improve the model fit at all and supports the conclusion that there is little evidence for year effects in the data. There was only weak support for the three-parameter nonlinear model that used date of entry as the independent variable (model 11), in part because it did not include gender as a covariate. Excluding the effect of gender, the nonlinear model still had lower support compared with its two-parameter equivalent (model 3) because there was no evidence of a nonlinear response of survey life to date of entry in the data. However, a linear model cannot be used in the escapement model without constraints or a transformation because it can predict a negative survey life for late-entry fish or departure dates that exceed the duration of the run for early-entry fish. The three-parameter nonlinear model had more support than the arctangent-transformed two-parameter one (model 4) and was therefore selected for use in the escapement model.

The limited survey life data ($n = 33$) suggests that the timing of entry into the survey area is equivalent for males and females but that males remain in the survey area longer. This dynamic was confirmed by the more extensive departure-timing data set ($n = 104$). The date at which 50% of the females had departed occurred 2 weeks earlier relative to males (Fig. 2*a*). The use of separate beta distributions predicting departure schedules for males and females provided a significant ($p = 2.03 \times 10^{-6}$) improvement in fit relative to a single beta distribution fit to all the data. Year-specific departure models also provided a significant improvement in fit ($p = 0.012$) relative to the single model. The maximum difference among years at which 50% of the spawners had departed (2003 vs. 2004) was only 11 days (Fig. 2*b*). Based

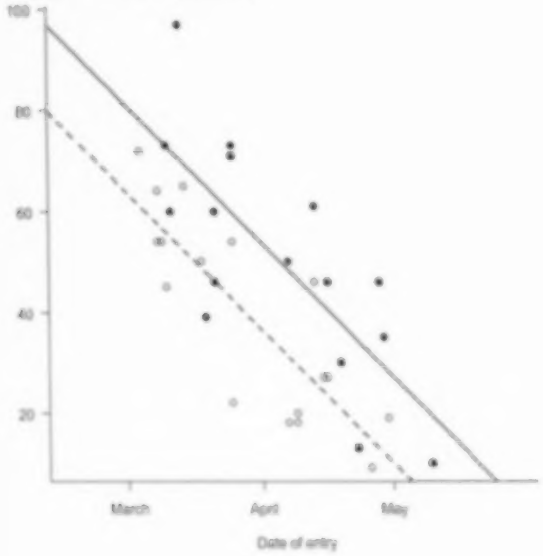
© 2007 NRC Canada

Table 2. Small sample size Akaike information criterion (AIC_c) statistics comparing out-of-sample predictive power for a range of models predicting survey life.

Model no.	Model form	No. of pars	r^2	Log-like	AIC_c	ΔAIC_c
1	B0	1	0.00	–64.32	130.77	10.40
2	M:B0, F:B1	2	0.07	–63.78	131.96	11.60
3	B0+B1*Day	2	0.50	–59.41	123.22	2.86
4	181*(Pi/2*ATan(B0 + B1*Day))/Pi	2	0.50	–61.67	127.74	7.38
5	B0+(M:B1*Day, F:B2*Day)	3	0.65	–56.86	120.55	0.19
6	**(M:B0, F:B1) + B2*Day**	**3**	**0.65**	**–56.77**	**120.36**	**0.00**
7	(M:B0 + B1*Day, F:B2 + B3*Day)	4	0.65	–56.76	122.96	2.60
8	($B0_{2003}$, $B1_{2003}$, $B2_{2004}$, $B3_{2005}$) + B4*Day	5	0.55	–58.65	129.53	9.17
9	B0 + B1*Day + B2*Yr	3	0.51	–59.26	125.34	4.98
10	B0 + (M:B1*Day, F:B2*Day) + B3*Yr	4	0.65	–56.70	122.83	2.47
11	B0*(1 – Day^B2/(B1^B2 + Day^B2))	3	0.50	–59.39	125.62	5.25

Note: *M* and *F* denote male and female, respectively. Log-like is the sum of log-likelihoods across all observations. The model with most support has the lowest AIC_c score and is highlighted in bold.

Fig. 1. Relationship between survey life and date of entry for males and females and best-fit models with gender-dependent intercepts but the same slope (Table 2, model 6). Solid and broken lines represent model predictions for males and females, respectively. Solid and open circles denote observations for males and females, respectively.



on AIC_c criteria, there was strong support for the gender-specific departure-timing model and no support for the gender-neutral (ΔAIC = 21.93) or year-specific (ΔAIC = 19.04) models.

Average catchability across the 52 surveys when tags were present was 0.22 and was highly variable (coefficient of variation (CV) = 0.74). The lower survey area is confined by dikes, has a lower gradient, and has less instream complexity relative to the upper survey area. As a result, average catchability was twofold higher (0.27) in the lower survey area than in the upper area (0.13). There was near strong support for the linear model predicting catchability based on HV/*Q* (Table 3, model 3). The ratio-based model provided substantially better fits relative to simpler models using only HV (model 1) or *Q* (model 2). The slightly lower ΔAIC value for the model that included a day-of-run effect (model 4) relative to the one that did not (model 3) suggests that catchability increases over the run when HV/*Q* effects are taken into account. Catchability increased with increases in the proportion of tags in the lower survey area (model 5) and water temperature (model 6), but there was only weak support for these models. There was essentially no support for models that included a year effect (models 7 and 8). The nonlinear three-parameter model that used HV/*Q* as an independent variable (model 9) had the strongest support for all models that were evaluated and was selected for use in the escapement model. This relationship explained 54% of the variation in catchability for surveys with nine or more tags present (n = 36).

Escapement, survey life, and timing estimates

An examination of the fit of the year-independent escapement model to the 2005 data (Fig. 3) is useful for characterizing its general behaviour. The model predicts relatively diffuse arrival timing (low values of τ) with a broad peak in the daily rate of arrival occurring between mid-March and early May (μ = 100; Fig. 3*a*). This arrival-timing results in a nearly linear increase in the cumulative number of arrivals over time (Fig. 3*b*). The predicted departure schedule, which depends on estimated arrival timing and the survey life – date of entry relationship (Fig. 3*d*), is consistent with observations (Fig. 3*b*) that show the vast majority of fish departing over the month of May. The escapement model assumes no process error, so parameters determining the total number of unmarked fish (U_i) are adjusted in the estimation to minimize the discrepancy between expanded counts (u_i/q_i) and predicted numbers present (Fig. 3*c*). The extent of adjustment can be assessed by comparing the prediction of the numbers present (Fig. 3*e*) with estimates based on simple expansions using the tag-recapture rate ($u_i/(r_i/R_i)$) or the qP–HV/*Q* relationship (u_i/qP_i) that assume no dependence of U_i over the run (Fig. 3*f*). The start of freshet in mid-April results in a decline in the ratio of horizontal visibility to dis-

© 2007 NRC Canada

Fig. 2. Cumulative proportion of fish exiting the survey area stratified by (*a*) gender and (*b*) year.

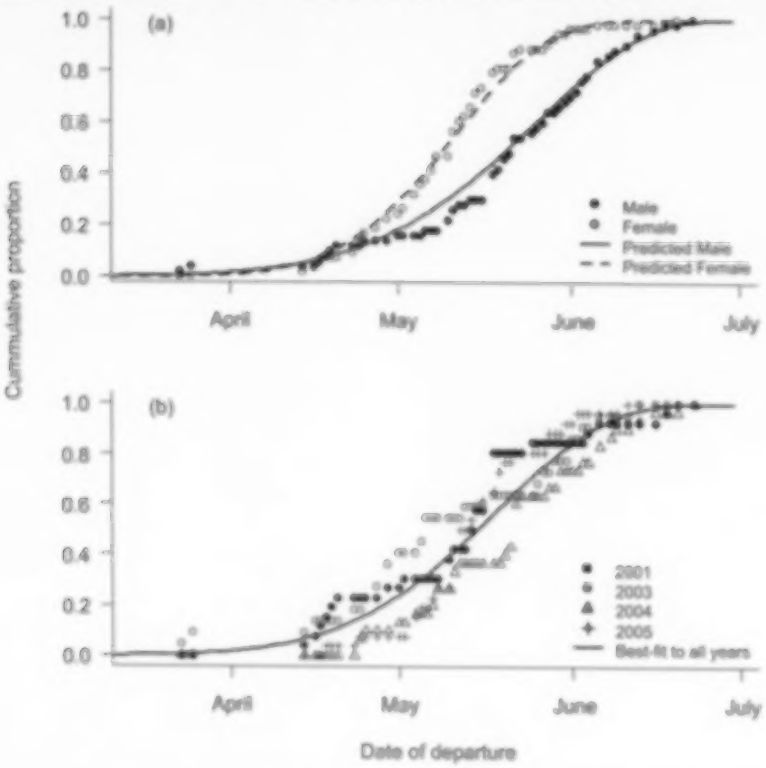


Table 3. Small sample size Akaike information criterion (AIC_c) statistics comparing out-of-sample predictive power for a range of models predicting observer efficiency.

Model no.	Model form	No. of pars	r^2 ($R_i > 8$)	Log-like	AIC_c	ΔAIC
1	B0 + B1*HV	2	0.34	−90.33	184.94	17.12
2	B0 + B1*Q	2	0.32	−87.40	179.09	11.28
3	B0 + B1*HV/Q	2	0.55	−83.21	170.70	2.88
4	B0 + B1*HV/Q + B2*Day	3	0.58	−81.78	170.15	2.34
5	B0 + B1*HV/Q + B2*Ptags	3	0.55	−83.20	172.99	5.18
6	B0 + B1*HV/Q + B2*Temp	3	0.59	−82.35	171.29	3.47
7	($B0_{2001}$, $B1_{2003}$, $B2_{2004}$, $B3_{2005}$) + B4*HV/Q	5	0.55	−82.85	177.25	9.43
8	B0 + ($B1_{2001}$, $B2_{2003}$, $B3_{2004}$, $B4_{2005}$)*HV/Q	5	0.54	−83.14	177.83	10.01
9	**(HV/Q)^B2/(B1^B2 + (HV/Q)^B2)**	**2**	**0.54**	**−81.76**	**167.81**	**0.00**

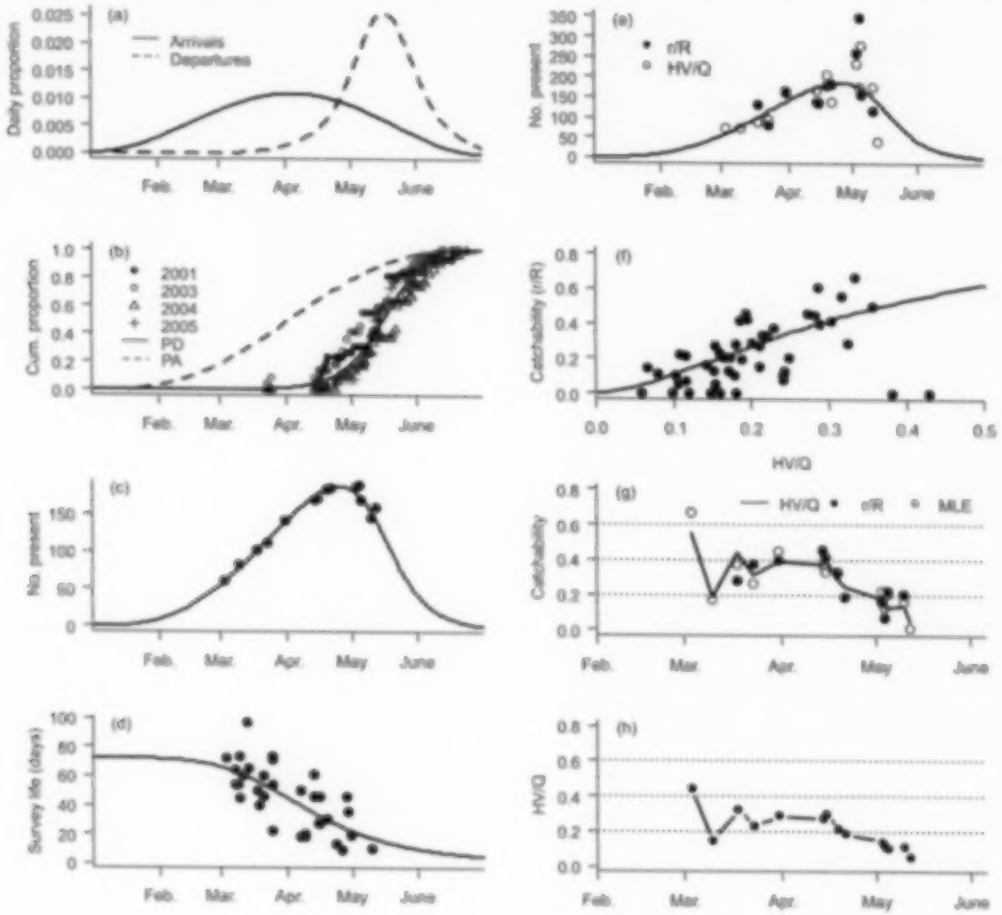
Note: HV, *Q*, Ptags, and Temp denote horizontal visibility, discharge, proportion of tags in the lower survey area, and water temperature, respectively. Log-like is the sum of log-likehoods across all observations. AIC_c values were computed based on all records in the data set for which information on all independent variables was available ($n = 45$ observations between 2001 and 2005). The r^2 values were computed for a subsample of the data where the number of tags present was >8 ($n = 36$). The model with most support has the lowest AIC_c score and is highlighted in bold.

charge (Fig. 3*b*) that in turn reduces catchability (Fig. 3*g*). Note that catchability remains consistently low after late April when the peak in the numbers present occurs. This reduces the contribution of the count data to the estimation of arrival-timing parameters.

Examination of the joint posterior distribution from the year-independent model for 2005 exposes the relationship among parameter estimates and the basis for model uncertainty (Fig. 4). There was a reasonably strong positive correlation between escapement and the mode of arrival timing; allowing more fish to enter late in the run when estimates of the numbers present are less certain because of reduced catchability produces higher escapement estimates. Survey life parameters were strongly correlated with each other because there are multiple ways for a three-parameter model to provide almost equivalent fits to data that can be well

© 2007 NRC Canada

Fig. 3. Predictions from the year-independent escapement model for 2005: (*a*) the proportion of the run arriving and departing by day; (*b*) the predicted cumulative (cum.) departure schedule for 2005 (PD), the observed departure schedules from 2001–2005, and the predicted cumulative arrivals (PA); (*c*) the predicted numbers present (line) and estimates (U_t) for each survey date (points); (*d*) the predicted survey life – date of entry relationship (line) and observations (points); (*e*) the predicted numbers present (line) and expanded counts (points) based on the survey-specific recapture rates (*r*/*R*) and river conditions (HV/*Q*); (*f*) predicted catchability (line) based on the ratio of horizontal visibility (HV) to discharge (*Q*) and estimates of catchability (*r*/*R*) from 2000–2005 (points); (*g*) the seasonal trend in catchability for 2005 based on either the conditional maximum-likelihood estimates (MLE), the *qP*–HV/*Q* relationship (HV/*Q*), or the recapture rate (*r*/*R*); and (*h*) the seasonal trend in the HV/*Q* ratio.

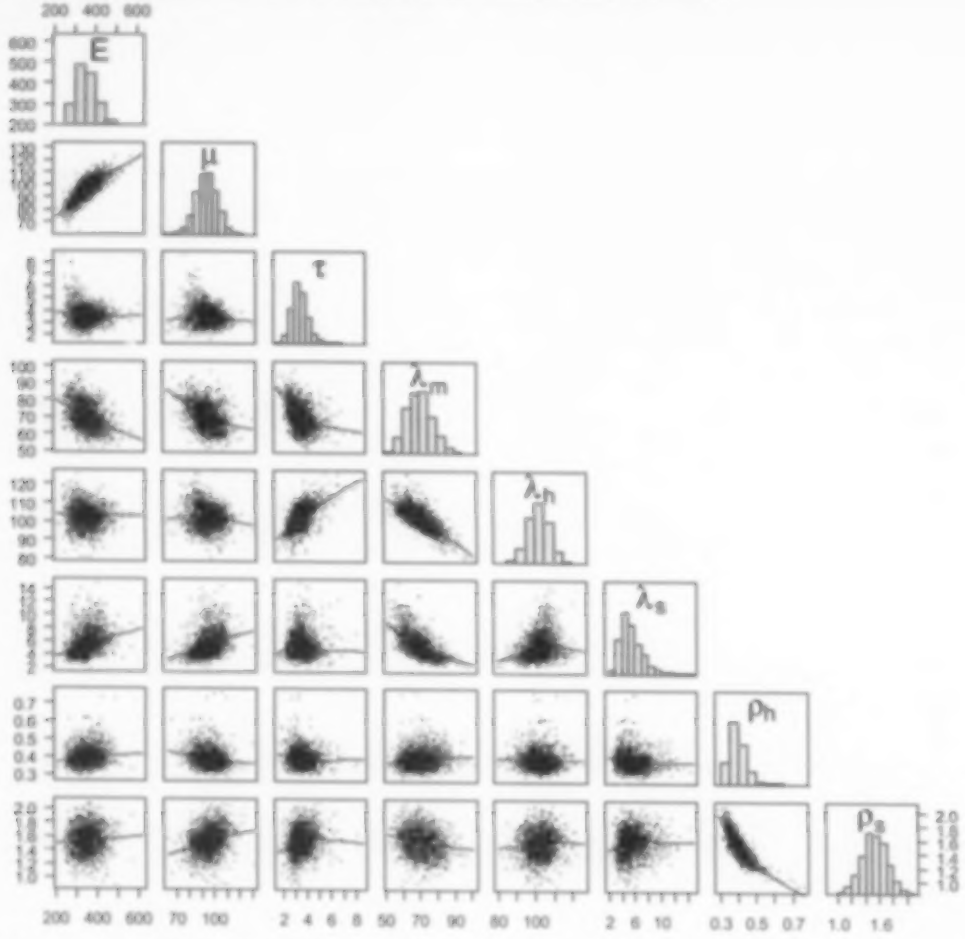


described by a simpler two-parameter linear model. This was highlighted by the AIC analysis (Table 2) in which the simple linear model (model 3) had much stronger support than the nonlinear model (model 11) used in the estimation. Low values of maximum mean survey life (λ_m) require higher values for λ_b and λ_s to fit the survey life data, with the net result of shorter survey life predictions, especially late in the run. As the number present is more uncertain late in the run than early in the run because of deteriorating river conditions, there is little penalty for parameters that suggest a greater proportion of spawners enter late in the run under the low survey life parameterization (note higher values of μ at low λ_m, and higher values τ at high λ_b), resulting in higher escapement estimates. More certain estimates of the numbers present late in the run combined with a well-defined peak across expanded counts would provide better definition of run-timing parameters and reduce the extent of confounding in survey life values, with a resulting increase in the precision of escapement estimates. There was a strong negative relationship in the *qP*–HV/*Q* parameters, which was not surprising as many catchability observations were highly uncertain and there is considerable variation around the best-fit relationship (Fig. 3*f*).

Escapement estimates between 1996 and 2006 (Table 4*a*) ranged from 90 to 469 based on the year-independent model (eq. 14) and from 117 to 429 when survey life and *qP*–HV/*Q* relationships were common among years (eq. 15). Average escapements over the 10-year time series were the

© 2007 NRC Canada

Fig. 4. Posterior distributions for the year-independent model parameter estimates for 2005 (histograms) and the correlation among parameters. A Lowess-smoothed trend line is used to highlight relationships among parameters.



same for both models, but there was more interannual variation associated with the year-independent model. Within-year coefficients of variation in escapement estimates for both models were typically higher between 1996 and 2000 when the number of surveys was low compared with later years. Escapement estimates from the common-year model had CVs that were on average 33% lower than those based on year-independent estimation. Within-year variation in escapement estimates based on the delta method were virtually identical to those calculated from the joint posterior distribution via MCMC simulation for the common-year model. CVs based on the delta method for the year-independent model were slightly lower than the MCMC-based CVs. There were asymmetries in confidence interval coverage for both models. We used delta-based CVs to examine the sensitivity of escapement estimates to various model assumptions and alternate data availability scenarios. However, confidence limits of escapement estimates should be calculated based on coverage probabilities defined by the MCMC analysis to account for asymmetries.

There was generally little variation in arrival and survey life parameters across years (Table 4*b*). Arrival-timing parameters were more variable under the year-independent model than the common model because the year-independent model allows for interannual variation in both the departure schedule and the qP–HV/Q relationship. The mode (μ) and precision of arrival timing (τ) ranged from 77 to 107 and from 2.8 to 4.9 for the year-independent model compared with 90 to 104 and 2.8 to 4.5 for the common-year model, respectively. Variability in survey life parameters was generally low and depended in part on the amount of information on run timing in the count data. For example, survey life predictions were lowest in 2002 because the count data in that year indicated a steep decline in the numbers present in May. The best way to fit all the data in this case was to predict late arrival timing (note higher values of μ and τ relative to other years) combined with reduced survey life. For comparison, we fit beta distributions to very accurate daily counts of winter-run steelhead moving upstream past a resistivity counter in the Keogh River, Vancouver Island, BC, between 1998 and 2005 (McCubbing 2002). Although the peak rate of arrival occurred a bit earlier at the Keogh River (mean μ = 85) relative to the Cheakamus (mean μ = 97), the precision of arrival timing was very simi-

© 2007 NRC Canada

Table 4. Year-independent (Ind.) and common-year (Com.) estimates of (*a*) escapement (*E*) and (*b*) arrival timing and survey life parameters.

(*a*) Escapement estimates.

	MLE of *E*		CV (delta) of *E*		CV (MCMC) of *E*	
Year	Ind.	Com.	Ind.	Com.	Ind.	Com.
1996 (4/0)	206	212	0.19	0.13	0.22	0.12
1997 (6/0)	109	125	0.22	0.14	0.24	0.14
1999 (5/0)	191	178	0.22	0.14	0.25	0.14
2000 (8/17)	90	117	0.22	0.15	0.22	0.15
2001 (9/27)	305	302	0.17	0.11	0.17	0.11
2002 (9/0)	469	429	0.16	0.10	0.18	0.10
2003 (16/33)	347	322	0.13	0.08	0.14	0.09
2004 (9/29)	351	346	0.16	0.11	0.19	0.11
2005 (14/37)	336	343	0.13	0.09	0.14	0.09
2006 (10/0)	255	281	0.16	0.10	0.17	0.10
Multiyear average	266	266	0.18	0.12	0.19	0.12
Multiyear CV	0.45	0.39				

(*b*) Estimates of arrival timing and survey life parameters.

	Arrival mode (μ)		Arrival precision (τ)		Survey life		
Year	Ind.	Com.	Ind.	Com.	λ_m	λ_b	λ_s
Common					69	103	4.61
1996	98	99	4.20	3.62	66	107	4.46
1997	89	92	3.10	3.53	76	99	4.13
1999	98	96	4.10	4.02	68	103	4.65
2000	77	93	2.10	2.82	86	94	3.85
2001	100	99	3.70	3.59	68	104	4.62
2002	107	104	4.90	4.52	60	108	6.25
2003	105	103	4.40	4.24	62	107	5.81
2004	104	103	4.10	3.64	67	104	4.21
2005	94	95	3.40	3.63	72	101	4.64
2006	86	90	2.80	3.05	78	98	4.03

Note: The coefficient of variation (CV) of escapement estimates was calculated using the Delta method and from the joint posterior distributions (MCMC). In (*a*), the total number of surveys per year and maximum tags in survey area are shown in parentheses after the year.

lar and very low (Keogh, τ = 1.6–3.9 with a mean of 2.3; Cheakamus, τ = 2.1–4.9 with a mean of 3.5).

Escapement model comparisons and sensitivity to data exchangeability assumptions

The common-year escapement model had much better out-of-sample predictive power relative to the year-independent model based on AIC_c and DIC criteria (Table 5). This is not surprising considering the extent of confounding between some model parameters seen in Fig. 4. Both ΔAIC_c and ΔDIC scores indicate that there was essentially no support for the year-independent escapement model. Similarly, there was a very high probability that the better fit of the year-independent model was due to chance alone based on a likelihood ratio test ($2\Delta L$ = 17.8, required ΔL at p = 0.05 = 61.7, $\chi^2_{2\Delta L\,df\,=\,80-35}$ = 1.000). The effective number of parameters (pD) computed for the DIC scores were 65 and 23 for year-independent and common-year models, respectively. pD values should approximate the actual number of parameters for nonhierarchical models when parameters are estimated using no constraints or informative prior information during the estimation (Gelman et al. 2004). Although we met these conditions for constraints and priors, there is an implicit hierarchical structure in both common-year and year-independent models that is imposed by using the same multiyear set of telemetry data in the estimation of year-specific escapement and arrival-timing values and because of the interaction among model parameters. The common-year model imposes more structure on the estimates by assuming that survey life and catchability relationships do not vary among years. As a result, the effective number of parameters is reduced by one-third relative to the number of true parameters (23 relative to a maximum of 35). The number of effective parameters for the year-independent model was reduced by only 18% (65 relative to a maximum of 80) and reflects the more flexible structure associated with allowing interannual variability in survey and catchability relationships. However, the higher DIC score for the year-independent model suggests that the increased flexibility is not supported by the data currently available.

We examined the sensitivity of escapement estimates to differences in data availability or assumptions about the exchangeability of run-timing data across years. Maximum-likelihood estimates of escapement and their standard deviations were computed across the range of likelihood structures described in Table 1 for both year-independent and

© 2007 NRC Canada

Table 5. Deviance information criterion (DIC) and second-order Akaike information criterion (AIC_c) statistics comparing out-of-sample predictive power of year-independent (eq. 14) and common-year (eq. 15) escapement models.

	Independent	Common
DIC		
D_{avg}	14 451	14 415
D_{exp}	14 386	14 392
pD ($D_{avg} - D_{exp}$)	65	23
DIC ($D_{avg} + pD$)	14 516	14 438
ΔDIC	78	
AIC		
Log-likelihood (L_f)	–7 181	–7 190
Deviance ($D = -L_f$)	14 362	14 379
No. of parameters (k)	80	35
AIC_c ($2D + 2k$·cf)	14 587	14 460
ΔAIC	127	

Note: D_{avg} and D_{exp} are the mean deviance of the posterior and the deviance of the posterior means, respectively; pD is the number of effective parameters; cf in the AIC_c calculation is the second-order small sample size correction factor computed as $n/(n - k - 1)$.

common-year models (Fig. 5). The comparison was done only for years with tagging data so the effect of not having year-specific tagging data could be evaluated, and results are shown using only 4 of 5 years of data for brevity. Excluding the year-specific tagging information (Table 1, likelihood 2) had little effect on escapement estimates because the tag-based survey-specific catchability estimates were generally highly uncertain. The year-independent model was more sensitive than the common-year model to excluding year-specific likelihoods determined by the qP–HV/Q relationship (likelihood 3), but the extent of this sensitivity depended on the year-specific information content of the tagging data (e.g., 2000 vs. 2005). The common-year model was much less sensitive to excluding L_{pu} and L_{pv} likelihood components because survey life and arrival parameters were more constrained. Excluding survey life (likelihood 4) followed a similar dynamic; variance in escapement estimates increased for the year-independent model only, and the extent of the effect depended on the year-specific information content. Most likely estimates of escapement increased substantially when departure data were excluded (likelihood 5) for both year-independent and common-year models because there was not enough information in the count or survey life data to discount the possibility of large numbers of fish entering late in the run. There was a very large increase in escapement estimates under the likelihood structure used in Korman et al. (2002), which excluded L_{pu}, L_{pv}, and departure data (likelihood 6) for both models in most years. Variance in escapement estimates also increased substantially, but this effect was variable across model types and years. Physically based predictions of catchability were more important in the estimation of the numbers present late in the run relative to early surveys because of reduced observer efficiency. In the absence of the qP–HV/Q model, the departure-timing data provided the only information to discount the probability of arrival-timing parameters that predicted a large number of fish entering late in the run, so removing both these components led to a large increase in escapement estimates. The extent of the differences in escapement and variance estimates depended on the information on run timing in the count data. Excluding year-specific tagging data was shown to have little influence on escapement estimates, so results for likelihood 7, which excluded this information and departure-timing data, were very similar to those of likelihood 5 in which only the departure data were excluded.

Discussion

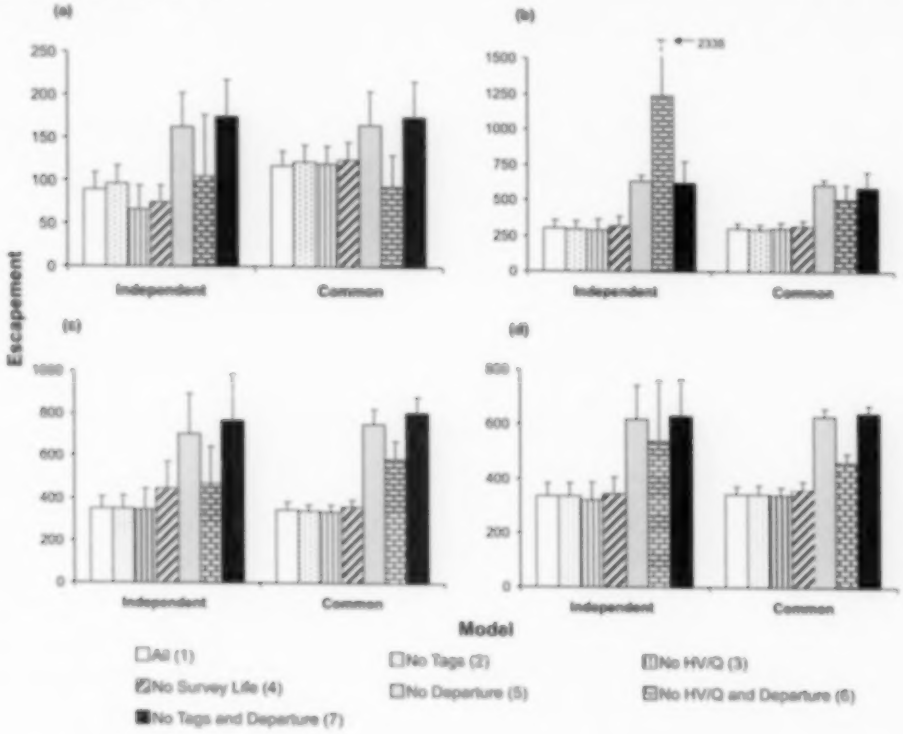
This analysis continues the advancement of maximum-likelihood and Bayesian models to estimate escapement from repeat-count data. The major advantage of our approach is that it increases the precision of escapement estimates and avoids the use of arbitrary constraints and overly simplistic assumptions about run timing. This is accomplished by jointly fitting run-timing parameters to count, survey life, and departure data and by reducing uncertainty in catchability by using a physical model that integrates mark–recapture estimates over multiple years. The relative variation in escapement estimates based on the year-independent model presented here were typically 40% less than those from a previous version of the year-independent model, which excluded the use of departure-timing data and the qP–HV/Q physical relationship (Korman et al. 2002). In contrast to previous efforts (Hilborn et al. 1999; Korman et al. 2002; Su et al. 2002), the multicomponent likelihood structure we have developed accounts for uncertainty in all model relationships.

Repeat counts of returning spawners will likely contain little information about run timing if there is large uncertainty about the numbers present at the peak or late in the run. The use of both survey life and departure-timing information obtained from radiotelemetry data was critical in reducing the uncertainty in escapement estimates in this situation. Gender-stratified survey life and departure-timing relationships were much less variable, and their use in the escapement model would increase the precision of escapement estimates. However, as we are unable to reliably determine the gender of fish during snorkel surveys, gender-stratified estimation is not possible. Entry into the survey area was similar across genders, but females had shorter survey lives and therefore left the survey area earlier compared with males. Females of migratory iteroparous species such as steelhead emigrate shortly after spawning to increase their probability of survival and future reproduction (Quinn 2005). Our data show that males remain in the survey area considerably longer than females, probably to increase their chances of spawning one or more times. These data do not suggest that the increased residence time for males is caused by earlier arrival timing, as shown for some Pacific salmon populations (Morbey 2000).

The relationship between catchability and the ratio of horizontal visibility to discharge was quite variable. A large part of this variability was caused by the considerable uncertainty in many of the survey-specific catchabilities estimates owing to low numbers of tags coupled with generally low catch-

© 2007 NRC Canada

Fig. 5. Most likely escapements in (*a*) 2000, (*b*) 2001, (*c*) 2004, and (*d*) 2005 from year-independent and common-year models estimated using all likelihood components (eq. 13), as well as subsets of the total likelihood (likelihood numbers correspond to those in Table 1). The total number of surveys per year and maximum tags in survey area were (*a*) 2000, 8/17; (*b*) 2001, 9/27; (*c*) 2004, 9/29; (*d*) 2005, 14/37. Error bars denote +1 SD.



ability. In the multivariate models in which the effects of river conditions were controlled through HV/*Q*, there was weak evidence that catchability increased over the duration of the run. It is likely that this increase was the result of behavioural changes associated with spawning. The date of first spawning in each survey year was recorded based on the presence of redds observed during snorkel surveys. Steelhead in the Cheakamus River typically initiated spawning by the second week of April when water temperatures exceeded 6.5–7 °C. Once spawning began, fish exhibited a much reduced flight response, were seen in pairs or larger groups, and were more commonly observed holding in pool tail-outs. Prior to spawning, fish were typically observed holding in the deepest and darkest portions at the upstream pool entrances or near undercut banks and woody debris. It is likely that the probability of detecting a fish is higher when they exhibit courtship and spawning behaviours and frequent the shallower, brighter, and more open environments that characterize pool tail-outs. Although these dynamics appeared obvious to swimmers, behavioural effects on catchability were only weakly supported by the data because of the relatively large uncertainty in catchability estimates coupled with strong temporal covariation among independent variables. Most observations occurred during either periods of good visibility and low discharge early in the run before spawning when most fish were in the lower survey area where catchability was higher or later in the run after spawning had begun when river conditions had deteriorated and a higher proportion of fish were in the upper survey area where catchability was lower. To tease-out these effects, it would be necessary to tag many more fish to stratify the catchability analysis by survey area to remove the spatial component of variation and to have periods of clear and low water at the same time that spawning is occurring to limit the extent of parameter confounding. Neither of these situations is likely to occur given the relatively low numbers of fish in the system, the difficulty in catching them to apply tags, and temperature-driven spawn timing and runoff patterns.

The use of the *qP*–HV/*Q* model improved the precision of escapement estimates considerably even though the relationship was poorly defined by the data. Survey-specific mark–recapture based estimates of the catchability were generally very imprecise and have the potential to produce substantive biases in estimates of the numbers present if the catchability of the tagged fish is not representative of untagged ones. In the case of Cheakamus River steelhead, overall catchability will be underestimated if the majority of tagged fish are in the upper survey area where catchability is lower and a large group of unmarked fish enter the river and hold in the lower survey area. If this occurs on a survey late in the run when there are few if any adjacent reliable estimates of the numbers present, the estimation model could substantially overestimate the escapement. Using the *qP*–HV/*Q* relationship

© 2007 NRC Canada

produces more robust estimates of survey-specific catchabilities compared with using tagging data for each survey independently because it averages over multiple surveys with similar river conditions, thereby improving precision and reducing the likelihood of large biases. Furthermore, the qP–HV/Q model allows count data to be used early and late in the run when there are no or very few marked fish present in the survey area in years with tagging and is the only way to estimate escapement in years when tagging is not conducted.

Survey life parameter estimates from the joint posterior distribution were quite confounded, and an escapement model based on a simpler relationship would likely have better numerical estimation characteristics. The confounding occurred because we used a three-parameter nonlinear survey life relationship rather than a two-parameter linear model that had the best out-of-sample predictive power based on the AIC_c criterion. When full uncertainty in all model parameters is admitted through the Bayesian implementation of the escapement model, the linear survey model would predict exit dates that exceed the assumed last day of the run and negative or unrealistically low survey lives for late-entry fish. The arctangent-transformed linear model would have avoided this problem, but its high AIC_c score showed that it did not fit the data very well. Even if there was substantial support for this model, we would be wary of using it because the true uncertainty in survey life during the early and late part of the run would be buried in the structure of the transformation. The nonlinear three-parameter model allows a wider range of survey life estimates during these periods and therefore admits more uncertainty in run timing. As there is little information on the relative abundance of fish late in the run from the count data, this results in an increase in the upper confidence limit of escapement estimates. The departure timing data, which does not support the hypothesis that there are large numbers of fish exiting very late in the run, reduces the probability of this dynamic, which in turn leads to reduced variation in escapement estimates and reasonable numerical performance. Our approach makes much more sense than using an unrealistic biological relationship that is supported solely from statistical criteria and better MCMC mixing properties.

The critical assumption in the model presented here is that data defining the relationships between catchability and river conditions, survey life and date of entry, and departure timing are completely exchangeable among years. This assumption appears reasonable based on both the frequentist and information theoretic model selection criteria that were applied to component relationships, as well as to the integrated escapement model. However, statistical criteria need to be applied cautiously when making decisions on model structure; the failure to detect year effects could be due to limited data rather than the absence of significant interannual variation in run-timing components. The extent of interannual variation in arrival and departure timing estimated by the model was in part limited because data from all telemetry years were pooled in the model-fitting process. This could result in an underestimate of the uncertainty in escapement estimates. The extent of this bias cannot be quantified with the data at hand, but we suspect that it is probably not large. Although differences in departure schedules among years were significant, they were separated by only 11 days in the most extreme case. It is probably reasonable to assume little variation in departure timing as very strong stabilizing selection on spawn and emergence timing has been documented in Atlantic salmon (*Salmo salar*; Jensen and Johnsen 1999; Einum and Fleming 2000; Letcher et al. 2004), steelhead (Hendry et al. 2002; Robards and Quinn 2002), and Pacific salmon (Quinn et al. 2000; Quinn 2005). As the departure schedule of kelting steelhead will be determined by spawn timing, strong stabilizing selection implies that the extent of interannual variation in departure timing will be limited. Both common-year and year-independent models predict that concentrated rates of arrival at the end of the run are unlikely. It is reassuring that this prediction is consistent with the arrival pattern of winter-run steelhead seen over many years at the Keogh River. However, additional measurements of survey life and departure timing for Cheakamus River spawners entering late in the season is required to confirm these run-timing hypotheses.

A logical next step in the analysis may be to develop a formal hierarchical model in which interannual variation in telemetry data is explicitly considered. Estimating a 10-year time series of escapement would require the joint estimation of hyperparameters for the interannual means and variances for all catchability and run-timing parameters ($k = 14$) in addition to 80 year-specific ones (total $k = 94$). Further, 28 values defining hyperpriors for the mean and variance of each hyperparameter would need to be specified. Considering the limited sample size of the survey life ($n = 33$), qP–HV/Q ($n = 52$), and departure timing data ($n = 104$) spread unevenly over 4–5 years, it is unlikely that such a complex model could be successfully applied without using reasonably informative priors for most of the parameters. Su et al. (2001) were not able to get the most complex of their relatively simple hierarchical models to converge in spite of 25 years of count data, a static and common-year survey life relationship, and the assumption that there was no error in the count data. Thus, although a formal hierarchical analysis of the Cheakamus steelhead data may be warranted in the future, it would be premature at this point as there is currently not enough information to accurately characterize the extent of interannual variation in catchability, survey life, and departure timing.

Acknowledgements

We thank the many anglers who captured fish for tagging. We also thank Brent Mossop, Howard Bailey, Don McCubbing, and two anonymous reviewers for providing helpful comments on an earlier draft and Steve Martell and Carl Walters for their guidance on Bayesian modelling. This work was funded by the British Columbia Hydro Power Authority and the Habitat Conservation Trust Fund of British Columbia.

References

Adkison, M.D., and Su, Z. 2001. A comparison of salmon escapement estimates using a hierarchical Bayesian approach versus separate maximum likelihood estimation of each year's return. Can. J. Fish. Aquat. Sci. **58**: 1663–1671.

© 2007 NRC Canada

Burnham, K.P., and Anderson, D.R. 2002. Model selection and multimodel inference. 2nd ed. Springer, New York.

Einum, S., and Fleming, I.A. 2000. Selection against late emergence and small offspring in Atlantic salmon (*Salmo salar*). Evolution, **54**: 628–639.

English, K.K., Bocking, R.C., and Irvine, J.R. 1992. A robust procedure for estimating salmon escapement based on the area-under-the-curve method. Can. J. Fish. Aquat. Sci. **59**: 1982–1989.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 2004. Bayesian data analysis. Chapman and Hall/CRC, Boca Raton, Fla.

Hendry, M.A., Wenburg, J.K., Myers, K.W., and Hendry, A.P. 2002. Genetic and phenotypic variation through the migratory season provides evidence for multiple populations of wild steelhead in the Dean River, British Columbia. Trans. Am. Fish. Soc. **131**: 418–434.

Hilborn, R., and Mangel, M. 1997. The ecological detective: confronting models with data. Monographs in Population Biology No. 28. Princeton University Press, Princeton, N.J.

Hilborn, R., Bue, B.G., and Sharr, S. 1999. Estimating spawning escapements from periodic counts: a comparison of methods. Can. J. Fish. Aquat. Sci. **56**: 888–896.

Jensen, A.J., and Johnsen, B.O. 1999. The functional relationship between peak spring floods and survival and growth of juvenile Atlantic salmon (*Salmo salar*) and brown trout (*Salmo trutta*). Funct. Ecol. **13**: 778–785.

Korman, J., Ahrens, R.N., Higgins, P.S., and Walters, C.J. 2002. Effects of observer efficiency, arrival timing, and survey life on estimates of escapement for steelhead trout (*Oncorhynchus mykiss*) derived from repeat mark–recapture experiments. Can. J. Fish. Aquat. Sci. **59**: 1116–1131.

Letcher, B.H., Dubreuil, T., O'Donnell, M.J., Obedzinski, M., Griswold, K., and Nislow, K.H. 2004. Long-term consequences of variation in timing and manner of fry introduction on juvenile Atlantic salmon (*Salmo salar*) growth, survival, and life-history expression. Can. J. Fish. Aquat. Sci. **61**: 2288–2301.

McCubbing, D. 2002. Adult steelhead trout and salmonid migration at the Keogh River, B.C., during spring 2002. Report prepared for BC Ministry of Water, Land, and Air Protection, Biodiversity Branch, Fisheries Research and Development available online at http://www.bccf.com/steelhead/pdf/Keogh_Salmonid_Enumeration_Report_ 2002.pdf

Morbey, Y. 2000. Protandry in Pacific salmon. Can. J. Fish. Aquat. Sci. **57**: 1252–1257.

Otter Research Ltd. 2004. An introduction to AD Model Builder version 7.1.1 for use in nonlinear modeling and statistics. Report available from otter-rsch.com.

Quinn, T.P. 2005. The behaviour and ecology of Pacific salmon and trout. American Fisheries Society, Bethesda, Md.

Quinn, T.P., Unwin, M.J., and Kinnison, M.T. 2000. Evolution of temporal isolation in the wild: genetic divernce in timing of migration and breeding by introduced Chinook salmon populations. Evolution, **54**: 1377–1385.

Robards, M.D., and Quinn, T.P. 2002. The migratory timing of adult summer-run steelhead in the Columbia River over six decades of environmental change. Trans. Am. Fish. Soc. **131**: 523–536.

Slaney, T.L., Hyatt, K.D., Northcote, T.G., and Fielden, R.J. 1996. Status of anadromous salmon and trout in British Columbia and Yukon. Fisheries, **21**: 20–35.

Su, Z., Adkison, M.D., and Van Alen, B.W. 2001. A hierarchical Bayesian model for estimating historical salmon escapement and escapement timing. Can. J. Fish. Aquat. Sci. **58**: 1648–1662.

Walters, C., and Ludwig, D. 1994. Calculation of Bayes posterior probability distributions for key population parameters. Can. J. Fish. Aquat. Sci. **51**: 713–722.

© 2007 NRC Canada

Geomorphology and stream habitat relationships with smallmouth bass (*Micropterus dolomieu*) abundance at multiple spatial scales in eastern Oklahoma

Daniel C. Dauwalter, Dale K. Splinter, William L. Fisher, and Richard A. Marston

Abstract: Fluvial geomorphic processes structure habitats important to stream fishes. We determined relationships between densities of smallmouth bass (*Micropterus dolomieu*) and ecoregions, watershed and reach morphology, and stream habitat in eastern Oklahoma, USA. Watershed and reach morphology were measured at 128 stream sites, and stream habitat and smallmouth bass abundance were measured in 1800 channel units. Variation in stream size, channel morphology, and substrate size constituted major physical differences among sites. Channel morphology differed among ecoregions in the largest streams. Densities of age-0 and age-1 and older smallmouth bass were approximately an order of magnitude greater in the Boston Mountains and Ozark Highlands streams than in Ouachita Mountains streams. Regression tree analysis explained less variation in age-0 (10-fold cross-validated relative error = 0.843) than in age-1 and older (relative error = 0.650) smallmouth bass densities and showed that stream size and channel-unit size were primary determinants of density. Channel morphology explained variation in densities in deep channel units of large streams, which was somewhat independent of ecoregion.

Résumé : Les mécanismes de géomorphologie fluviale structurent les habitats qui sont importants pour les poissons d'eau courante. Nous avons déterminé les relations entre les densités d'achigans à petite bouche (*Micropterus dolomieu*) et les écorégions, la morphologie du bassin versant et des sections, ainsi que les habitats des cours d'eau dans l'est de l'Oklahoma, É.-U. Nous avons mesuré la morphologie du bassin versant et de la section à 128 sites et déterminé l'habitat fluvial et la densité des achigans à petite bouche dans 1800 unités de chenal. Les variations de taille du cours d'eau, de la morphologie du chenal et de la granulométrie du substrat constituent les différences physiques majeures entre les sites. La morphologie du chenal varie dans les plus grands cours d'eau d'une écorégion à une autre. Les densités des achigans d'âge 0, d'âge 1 et d'âges supérieurs sont environ 10 fois plus importantes dans les cours d'eau des monts Boston et des terres hautes de l'Ozark que dans ceux des monts Ouachita. Une analyse par arbre de régression explique moins de variation (erreur relative par validation croisée en 10 parties = 0,843) de la densité des achigans à petite bouche d'âge 0 que de la densité des poissons d'âge 1 ou plus (erreur relative = 0,650); elle montre aussi que la taille du cours d'eau et celle de l'unité de chenal sont les principaux facteurs déterminants de la densité. La morphologie du chenal explique les variations de densité dans les unités de chenal profondes des grands cours d'eau, en partie indépendamment de l'écorégion.

[Traduit par la Rédaction]

Introduction

The processes of water, sediment, and wood transport control stream channel morphology and hydraulics and ultimately form the physical habitat in streams required by fishes (Yang 1971; Beschta and Platts 1986; Newson and Newson 2000). Downstream transport of water, sediment, wood, and nutrients creates the longitudinal profile of streams that have unique characteristics, processes, and biota (Vannote et al. 1980). However, spatial variability in those routing processes within and among watersheds and regions exists as a result of landforms, geomorphic controls, and dis-

Received 9 May 2006. Accepted 24 March 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 9 August 2007. J19312

D.C. Dauwalter,[1,2] Oklahoma Cooperative Fish and Wildlife Research Unit and Department of Zoology, Oklahoma State University, Stillwater, OK 74078, USA.
D.K. Splinter. Department of Geography and Geology, University of Wisconsin–Whitewater, WI 53190, USA.
W.L. Fisher. United States Geological Survey, Oklahoma Cooperative Fish and Wildlife Research Unit and Department of Zoology, Oklahoma State University, Stillwater, OK 74078, USA.
R.A. Marston. Department of Geography, Kansas State University, Manhattan, KS 66506, USA.

[1]Corresponding author (e-mail: ddauwalt@uwyo.edu).
[2]Present address: Department of Zoology and Physiology, Department 3166, University of Wyoming, 1000 East University Avenue, Laramie, WY 82071, USA.

 doi:10.1139/F07-085 © 2007 NRC Canada

turbances that can affect local channel morphology and stream habitat within the longitudinal continuum (Montgomery 1999; Benda et al. 2004).

Channel morphology and landforms affect stream biota at different organismal and organizational levels. In an Appalachian stream, channel morphology and geomorphic controls (bedrock outcrops) influenced macroinvertebrate community structure and function (Huryn and Wallace 1987). Bull trout (*Salvelinus confluentus*) selected reaches for spawning that were in alluvial valleys with geomorphic controls (knickpoints) and large areas of groundwater upwelling. Within those reaches, however, bull trout built redds at the transitions of bedforms that had areas with localized downwelling (Baxter and Hauer 2000). Watershed and reach morphology have been shown to structure stream fish assemblages (Rhoads et al. 2003; Walters et al. 2003), with the importance of spatial scale being dependent on the spatial extent of the study (Gido et al. 2006). Fluvial processes acting on different spatial scales have been implicated in affecting stream habitat and biota; however, connecting these processes to stream morphology, physical habitat, and biota in streams often requires incorporation of analyses at multiple spatial scales (Frissell et al. 1986; Newson and Newson 2000; Frothingham et al. 2002).

Many stream habitat enhancement and restoration principles are based on the characteristics of watersheds, valleys, and stream channels. Rosgen (1994, 1996) developed a stream classification system to determine, in part, the potential for enhancement of fish habitat and river restoration. He suggested that installation of instream structures intended as fish habitat may result in detrimental channel adjustments and create disequilibrium conditions and that only streams exhibiting certain morphologic and sediment characteristics should be candidates for habitat-improvement structures. Rosgen also suggested that restoration principles should incorporate the natural stable tendencies of a river that result from the interaction of morphological variables. Champoux et al. (2003) found that long-term success of instream structures was related to different glacial deposits (moraine versus outwash plain) and argued that the glacial history and reach characteristics need to be considered during restoration planning to ensure long-term effectiveness of habitat-improvement structures and management actions.

Because stream restoration and enhancement activities incorporate the geomorphic context of streams at several spatial scales, understanding how fishes are related to the geomorphic features of streams aids the conservation and management of local stream habitats and populations of aquatic organisms. Our goal was to determine how watershed characteristics, channel morphology, and stream habitat influenced smallmouth bass (*Micropterus dolomieu*) abundance in streams of the Boston Mountains, Ouachita Mountains, and Ozark Highlands ecoregions in eastern Oklahoma. These ecoregions in Oklahoma encompass the southwestern extent of the native distribution of smallmouth bass.

Materials and methods

Study area

Relations between smallmouth bass abundance and watershed characteristics, channel morphology, and stream habitat were determined for small streams to midsized rivers in the Boston Mountains, Ouachita Mountains, and Ozark Highlands level-III ecoregions in eastern Oklahoma (Omernik 1987). The Boston Mountains ecoregion is a dissected mountainous plateau composed of flat-lying shale and sandstone lithology and oak–hickory forest. Annual rainfall is 112–130 cm and land is used mostly for logging and recreation (Woods et al. 2005). The Ouachita Mountains consist of folded, Paleozoic sandstone, shale, and chert. Oak – hickory – shortleaf pine forests in this region receive from 109 to 145 cm of annual rainfall. Common land uses are logging, pastureland, hayfields, farming, and recreation (Rutherford et al. 1992; Woods et al. 2005). The Ozark Highlands are dominated by flat-lying cherty limestone, but shale, limestone, and dolomite are present in valley bottoms. Annual precipitation is 104–124 cm, and land use is primarily logging, recreation, and cattle and poultry farming, including pastureland and hayfields (Rabeni and Jacobson 1993; Woods et al. 2005).

Stream survey

A geographic information system (GIS) was used as an aid in the random selection of 175 stream sites for an inventory of watershed and reach morphology, stream habitat, and smallmouth bass populations in the study area. To select stream sites, a stream network was generated in a GIS using a 30 m digital elevation model (DEM) from the National Elevation Dataset (US Geological Survey 1999). The network was created by identifying all cells in the DEM with a watershed size ≥1.35 km^2 and defining them as a stream. Stream orders (Strahler 1957) were assigned to each stream segment in the network. One-hundred and seventy-five sample sites (sample units) were randomly selected on regional streams. Sites were allocated to the three ecoregions approximately in proportion to ecoregion area and equally distributed among stream orders 1–4 in each ecoregion to ensure that larger streams were represented. Forty sites were selected in both the Boston Mountains and Ozark Highlands ecoregions, and 95 were selected in the Ouachita Mountains ecoregion within Oklahoma. The GIS layer of sample sites was uploaded into a global positioning system (GPS) receiver that was used to navigate to each site. Streams were sampled once from mid-May to mid-August from 2003 to 2005 during low streamflow conditions.

Watershed and channel morphology

The characteristics of watersheds and stream channels associated with each stream site were measured. Watershed scale geomorphic variables for each site were derived using a GIS. The aforementioned DEM was used to delineate watersheds, and GIS was then used to measure watershed areas and lengths. Elongation ratio was measured as the diameter of a circle with the same area as the watershed divided by watershed length (Morisawa 1968).

We measured channel morphology at the reach scale. The stream reach at each site was defined as 20 times the mean channel width (i.e., bankfull width) upstream of each sample point; bankfull stage was identified by using visual and physical indicators (Rosgen 1996). Channel units in each reach were classified by relative water velocities and surface turbulence, using the scheme of Hawkins et al. (1993). Transects perpendicular to the channel were surveyed across

© 2007 NRC Canada

two riffles and two pools when available, with a maximum of two transects in a single channel unit; two to four transects were surveyed per reach. Entrenchment ratio, width–depth ratio, and median particle size (i.e., D_{50}) were calculated for each transect. Entrenchment ratio was calculated as the ratio of flood-prone width to bankfull channel width. Flood-prone elevation was equal to twice the maximum channel depth, and flood-prone width was estimated as the valley floor width at the flood-prone elevation. Maximum flood-prone width was measured if it was less than 2.2 times the channel width at each transect, otherwise it was recorded as 2.2. Width–depth ratio was calculated as the ratio of bankfull channel width to mean bankfull channel depth. Channel depths were measured at 20 equally spaced locations along transects. Median particle size of surficial substrates was calculated by collecting 100 particles along each transect and measuring the intermediate axis of each; bedrock was excluded from analysis. Water surface slope of each reach was measured in the field with a transit level and stadia rod. Sinuosity was measured as the ratio of stream length to straight-line distance as measured on maps of the stream reach created with a GPS (see below). Width–depth ratio and median particle size for the reach were calculated from transect data as weighted averages based on the proportion of each reach length that comprised riffles and pools (*Zar* 1999). Reach entrenchment ratios were calculated as the median of transect values.

Channel-unit habitat

Habitat was measured in each channel unit within each reach using methods conducive to a large-scale stream survey. Channel units were mapped with a GPS, and lengths and areas of channel units were measured in a GIS (Dauwalter et al. 2006). Thalweg depths were measured multiple times per channel unit at equally spaced intervals; more measurements were made in longer and more heterogeneous channel units. A modified Wentworth scale was used to visually estimate substrate distributions within each channel unit (Bain 1999). We estimated, and enumerated when logistically feasible, boulders, rootwads, and large woody debris (10+ cm diameter, 4+ m in length) in each channel unit. Percent undercut bank and percent vegetative coverage were also estimated. Water velocity was coded by channel-unit type (falls, 4; riffle, 3; run, 2; all other types, 1). Habitat variables for each channel unit were combined for reach habitat estimates.

Smallmouth bass abundance

Smallmouth bass abundance was estimated by snorkeling or electrofishing individual channel units within reaches. When snorkeling, 1- to 3-person snorkel teams moved upstream in a zigzag pattern and counted and estimated sizes of smallmouth bass individuals (Dolloff et al. 1996). One electrofishing pass was made throughout each channel unit using a backpack, barge, or boat electrofisher when water clarity was less than 1 m and prohibited effective snorkeling. Water clarity was measured as the maximum distance at which a snorkeler could see a 100 mm model of a smallmouth bass. Electrofishing power density was standardized at approximately 1000 $\mu S \cdot cm^{-1}$ (Miranda and Dolan 2003); however, settings on the backpack electrofisher prohibited obtaining 1000 $\mu S \cdot cm^{-1}$ exactly. Counted individuals were placed into <100 and ≥100 mm size categories. Individuals < 100 mm represent age-0 individuals and those ≥ 100 mm are age-1 and older (Balkenbush and Fisher 2001); ages will be referred to hereafter.

Counts of smallmouth bass made in channel units while snorkeling and electrofishing were adjusted to estimate abundance. To adjust snorkel counts, we constructed a model using counts and habitat variables to predict abundance. The model was built with study-independent data collected from channel units that were snorkeled first and then electrofished with three passes. Electrofishing abundance estimates were made using a three-pass capture probability model in which the number of individuals collected was divided by electrofishing capture probability (Dauwalter and Fisher 2007*a*). Electrofishing abundance estimates were then modeled as a function of snorkel counts and count interactions with fish size (0, <100; 1, ≥100 mm) and habitat variables (water clarity, channel-unit depth, percent bedrock, and wood density). A no-intercept model was fit with all variables using SAS version 9.1 software (SAS Institute Inc., Cary, North Carolina). Variables with a significant effect ($P \leq 0.15$) on abundance estimates were retained and the model was refit. The final model used to adjust snorkel counts was

$$\hat{N} = 2.6176(\text{snorkel count}) + 0.1478(\text{snorkel count} \times \text{percent bedrock}) - 1.6354(\text{snorkel count} \times \text{size class})$$

The model fit well (adjusted $R^2 = 0.79$, $N = 59$). When channel units were electrofished because they were too turbid to snorkel, counts of smallmouth bass were adjusted using a one-pass electrofishing capture probability model (Dauwalter and Fisher 2007*a*).

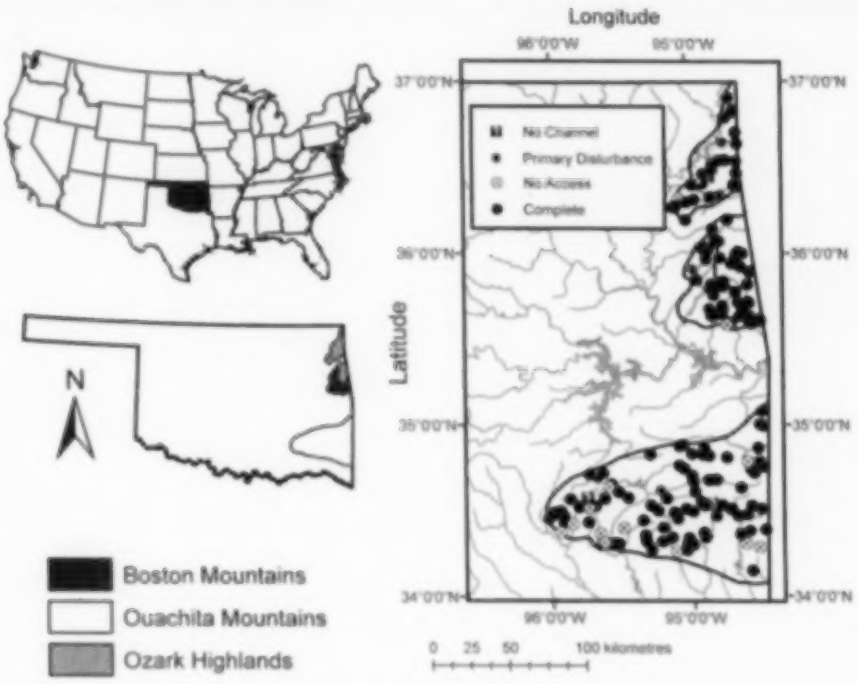
Data analysis

We used a principle components analysis (PCA) to determine relationships among variables at the watershed and reach scales and to eliminate redundant covariables. The PCA was conducted on the correlation matrix, and eigenvalues greater than those predicted under the broken-stick model were considered meaningful (Jackson 1993; McGarigal et al. 2000). The predicted eigenvalue (b) for the kth component under the broken-stick model was calculated as $b_k = \sum_{i=k}^{p} 1/i$, where p was the number of variables. Pearson correlations were used to examine relationships among channel-unit variables; the limited number of variables precluded use of PCA at the channel-unit spatial scale. Analyses were run using SAS version 9.1 software (SAS Institute Inc., Cary, North Carolina).

We used regression tree analysis (RTA) to determine the influence of ecoregions, watershed and reach morphology, and stream habitat at the reach and channel-unit spatial scales on the density of smallmouth bass in channel units. RTA is a flexible and robust nonparametric method used to assess complex relationships between explanatory variables and a response variable that may be nonlinear with high-order interactions. It explains variation in the response variable by splitting the data set into homogenous groups using

© 2007 NRC Canada

Fig. 1. Sites selected (175) and sampled (128) for an inventory of watershed and reach morphology, stream habitat, and smallmouth bass (*Micropterus dolomieu*) in eastern Oklahoma streams.



specific values of the categorical or continuous explanatory variables. Groups include observations that have common values of explanatory and response variables. Regression trees are displayed graphically for ease of interpretation (Breiman et al. 1984; De'ath and Fabricius 2000). RTA has been used in other ecological studies of complex relationships between explanatory and response variables when common linear modeling approaches have failed (Rejwan et al. 1999; De'ath and Fabricius 2000).

We conducted RTA for data collected at the ecoregion, watershed and reach, and channel-unit scales in one analysis for each age group. Although sample reaches were our defined sample units for the study, we defined each channel unit as an observation during RTA. Because RTA is a nonparametric analysis, the use of channel units as an observation when they are nested within reaches is not problematic. Explanatory variables included in RTA were dependent on results of PCA and correlation analyses to limit variable redundancy. Smallmouth bass densities (number·ha^{-1}) were natural log transformed ($\log(X + 0.01)$) to reduce the effects of zero densities on variances. Each RTA was performed with CART version 6.2 software (Steinberg and Colla 1995). We used sums of squares about the group means as the measure to maximize homogeneity within groups; sums of absolute deviations about the median are more robust but can be ineffective when many zeros are present (De'ath and Fabricius 2000). The final tree selected was the one with the smallest relative error (similar to the coefficient of nondetermination, $1 - r^2$) based on 10-fold cross-validation (Breiman et al. 1984; De'ath and Fabricius 2000).

Results

Stream survey

We surveyed channel morphology, stream habitat, and smallmouth bass abundance at 128 of 175 selected stream sites. Seventeen stream sites were either inaccessible or access was denied by landowners. Direct channel alteration by humans (e.g., gravel mining) was observed at five stream sites, and these sites were not sampled. Twenty-two streams were dry. Two sites were not snorkeled, and habitat measurements were incomplete at another. Complete information on watershed and reach morphology, stream habitat, and smallmouth bass abundance was collected at 128 stream sites (Fig. 1); electrofishing was conducted at 17 sites in the Ouachita Mountains. Among all sites, 1800 channel units were sampled.

Watershed and channel morphology and habitat

Watershed and channel morphology and stream habitat were more similar in the Boston Mountains and Ozark Highlands than in the Ouachita Mountains (Table 1). Ouachita Mountains streams had smaller watershed sizes on average, likely because most small streams in the other ecoregions were dry and excluded from analyses. Ouachita Mountains streams generally had higher reach slopes, larger particle sizes, and more vegetation, silt–clay substrate, and pool area. They were also typically deeper and had less wood and lower width–depth ratios than streams in the other regions. Stream reaches in the Boston Mountains were typically shallower and had more bedrock. Ozark Highland streams were

© 2007 NRC Canada

Table 1. Morphology of watersheds and stream reaches in the Boston Mountains (n = 29), Ouachita Mountains (n = 78), and Ozark Highlands (n = 21) ecoregions in eastern Oklahoma.

Variable		Boston Mountains	Ouachita Mountains	Ozark Highlands
Watershed area (km^2)	Mean	116.1	61.5	182.0
	Minimum	1.3	1.2	4.8
	Maximum	587.5	393.5	951.3
D_{50} (mm)	Mean	35.4	79.6	27.6
	Minimum	2.7	1.5	16.8
	Maximum	83.5	309.0	51.5
Elongation ratio	Mean	0.57	0.53	0.49
	Minimum	0.43	0.22	0.32
	Maximum	0.82	0.84	0.67
Large woody debris (no.·m^{-2})	Mean	0.0034	0.0030	0.0030
	Minimum	0.0000	0.0000	0.0000
	Maximum	0.0405	0.0235	0.0113
Mean thalweg depth (m)	Mean	0.30	0.37	0.43
	Minimum	0.11	0.10	0.22
	Maximum	0.72	1.89	0.83
Percent bedrock	Mean	16.9	10.5	8.4
	Minimum	0.0	0.0	0.0
	Maximum	88.5	67.7	68.8
Percent pool	Mean	63.8	86.2	62.9
	Minimum	11.6	23.2	34.6
	Maximum	98.9	100.0	100.0
Percent silt–clay	Mean	4.2	7.5	5.2
	Minimum	0.0	0.0	0.0
	Maximum	30.2	45.0	24.9
Percent vegetation	Mean	2.0	8.2	3.8
	Minimum	0.0	0.0	0.0
	Maximum	17.6	61.1	17.4
Rootwads (no.·m^{-2})	Mean	0.0033	0.0055	0.0072
	Minimum	0.0000	0.0000	0.0000
	Maximum	0.0595	0.0235	0.0593
Sinuosity	Mean	1.14	1.16	1.18
	Minimum	1.00	1.00	1.00
	Maximum	1.51	3.04	1.80
Slope	Mean	0.0069	0.0075	0.0045
	Minimum	0.0008	0.0000	0.0005
	Maximum	0.0274	0.0315	0.0194
Width–depth ratio	Mean	23.1	16.9	27.5
	Minimum	11.4	6.9	6.6
	Maximum	44.0	36.7	54.7
Wood (no.·m^{-2})	Mean	0.0067	0.0085	0.0102
	Minimum	0.0000	0.0000	0.0000
	Maximum	0.1000	0.1192	0.0593

deeper with more wood, wider channels, smaller particle sizes, and less bedrock on average when compared with streams in the other two regions. Variability in watershed and reach morphology and stream habitat was often greater in the Ouachita Mountains, the largest ecoregion.

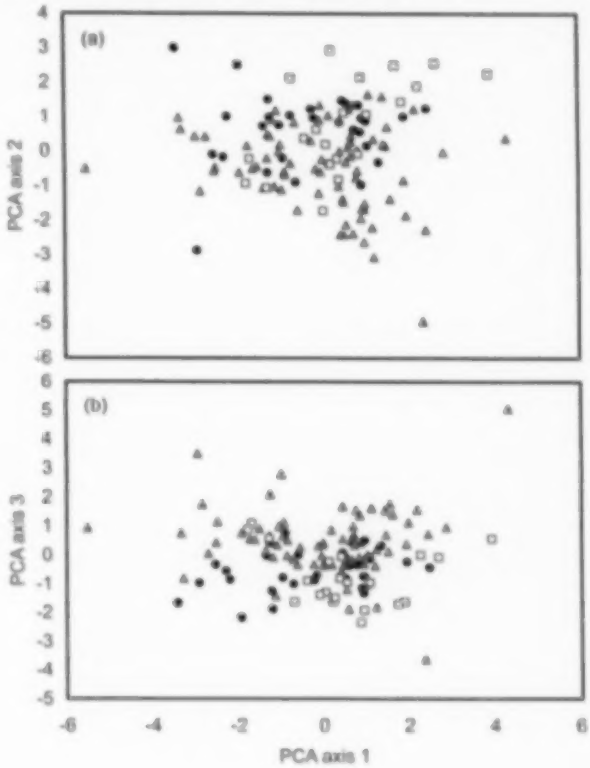
At the watershed and reach scales, PCA revealed two informative principle components according to the broken-stick criterion; however, the third principle component also appeared important (Table 2). Axis 1 reflected stream size. Larger streams were deeper and had lower channel slopes, confirming the expected relationships among these variables. Axis 2 represented channel morphology. Wide, shallow channels had less silts and clays, more riffle–pool structure, and more bedrock. Axis 3 represented a particle size axis, dominated by D_{50}, and followed by percent pool and sinuosity. PCA biplots showed that channel morphology (axis 2) did not differ in small streams (axis 1) among ecoregions, but it differed between the largest Ozark Highlands and Ouachita Mountains streams; channel morphology in the Ouachita Mountains was also more variable than in the Boston Mountains and Ozark Highlands (Fig. 2). The pattern of site scores on axis 3 suggested that streams in the Ouachita Mountains typically had larger substrates and more pool habitat and were less sinuous.

© 2007 NRC Canada

Table 2. Statistics and eigenvalues for eigenvectors from a principal component analysis (PCA) of geomorphic and stream habitat variables from 128 stream reaches in the Boston Mountains, Ouachita Mountains, and Ozark Highlands ecoregions in eastern Oklahoma.

	PC 1	PC 2	PC 3
Summary statistics			
Observed eigenvalue	2.448	1.785	1.345
Broken-stick eigenvalue	1.000	1.500	1.833
Percent of variance explained	20.4	14.9	11.2
Variables			
Watershed area (km^2)	0.402	0.332	0.055
D_{50} (mm)	–0.027	0.166	0.621
Elongation ratio	–0.255	0.018	0.273
Mean thalweg depth (m)	0.458	0.097	0.324
Percent bedrock	–0.086	0.353	–0.141
Percent pool	0.253	–0.349	0.427
Percent silt–clay	0.219	–0.447	–0.144
Percent vegetation	0.138	–0.113	0.009
Sinuosity	0.154	–0.277	–0.377
Slope	–0.544	0.020	0.115
Width–depth ratio	0.107	0.528	–0.208
Wood (no.·m^{-2})	–0.306	–0.195	0.100

Fig. 2. Principal component analysis (PCA) biplots of watershed and reach morphology and habitat of 128 stream reaches in the Boston Mountains (solid circles), Ouachita Mountains (shaded triangles), and Ozark Highlands (open squares) ecoregions, eastern Oklahoma. Axis 1 represented stream size, where slope (–0.544), mean thalweg depth (0.458), and watershed area (0.402) had the highest axis loadings (*a* and *b*). Axis 2 represented channel stability; width–depth ratio (0.528), percent silt–clay (–0.447), and percent bedrock (0.353) had high axis loadings (*a*). D_{50} (0.621), percent pool (0.427), and sinuosity (–0.377) had high axis loadings for axis 3 (*b*).



Channel-unit habitat

We sampled 11 channel-unit types in all. Riffles, runs, and midchannel pools were the most abundant types in each ecoregion. Habitat characteristics were highly variable among channel-unit types and ecoregions (Table 3). No variables among the 1800 channel units were highly correlated (Table 4). Velocity and mean thalweg depth had the highest Pearson correlation at $r = -0.398$.

Smallmouth bass abundance

Smallmouth bass densities varied widely but were typically higher in the Boston Mountains and Ozark Highlands. No smallmouth bass were collected in turbid stream reaches that were electrofished. Density estimates of age-0 smallmouth bass in channel units ranged from 0·ha^{-1} in all ecoregions to 9406·ha^{-1} in the Boston Mountains, 2331·ha^{-1} in the Ouachita Mountains, and 4873·ha^{-1} in the Ozark Highlands. Densities of age-1 and older smallmouth bass ranged from 0·ha^{-1} in all ecoregions to 8914·ha^{-1} in the Boston Mountains, 775·ha^{-1} in the Ouachita Mountains, and 4642·ha^{-1} in the Ozark Highlands. Densities were consistently high in runs and pools, although some riffles, abandoned channels, and backwaters also had high densities (Table 5).

Stream size, channel-unit depth, and reach morphology variables explained variation in densities of age-0 smallmouth bass. Seven watershed–reach and seven channel-unit scale variables were included in RTAs. Although reach slope was correlated with watershed area, we retained both because of the variability observed in the Ouachita Mountains ecoregion. Rootwad and large woody debris densities were combined (added) to form a wood density variable. Elongation ratio was not included because of the lack of direct biological influence on fish abundance. The regression tree with the minimum cross-validated relative error, 0.843, had five nodes, three that were terminal (Fig. 3). We conducted an exploratory analysis by extending the tree to nine and 15 nodes that had relative errors of 0.866 and 0.867, respectively. Thus, although the total amount of explained variation in densities of age-0 smallmouth bass was low (13%–15%), some variables explained variation at a higher probability than could be expected by chance. Age-0 densities were higher in deeper channel units in larger streams. In deeper channel units of larger streams, densities were higher in reaches with channel slopes greater than 0.07% and width–depth ratios less than 25. When width–depth ratios were higher than 25, age-0 densities were higher in deeper channel units. When densities were high in shallow channel units, they were high in higher gradient reaches (Fig. 3).

Several watershed–reach and channel-unit variables explained most of the variation in densities of age-1 and older

© 2007 NRC Canada

Table 3. Habitat characteristics of common channel-units types in streams of the Boston Mountains, Ouachita Mountains, and Ozark Highlands ecoregions in eastern Oklahoma.

Variable		Riffle	Run	Midchannel pool	Lateral pool	Backwater
Boston Mountains						
Number (*n*)		191	183	160	16	31
Mean depth (m)	Mean	0.156	0.307	0.329	0.526	0.304
	Minimum	0.03	0.07	0.05	0.07	0.07
	Maximum	0.55	1.08	1.67	1.04	0.72
Percent bedrock	Mean	19.4	23.9	20.3	4.4	4.2
	Minimum	0	0	0	0	0
	Maximum	100	100	95	20	70
Percent boulder	Mean	7.7	5.0	5.9	5.4	2.4
	Minimum	0	0	0	0	0
	Maximum	75	80	60	15	15
Percent silt–clay	Mean	0.4	1.3	2.9	12.3	10.1
	Minimum	0	0	0	0	0
	Maximum	10	20	80	85	50
Percent vegetation	Mean	2.6	1.5	1.4	1.8	3.9
	Minimum	0	0	0	0	0
	Maximum	80	70	35	10	60
Velocity (coded)	Mean	3	2	1	1	1
	Minimum	3	2	1	1	1
	Maximum	3	2	1	1	1
Wood (no.·m^{-2})	Mean	0.00875	0.01000	0.11812	0.08463	0.02307
	Minimum	0.0000	0.0000	0.0000	0.0000	0.0000
	Maximum	0.5882	0.6250	12.0482	0.9091	0.1427
Ouachita Mountains						
Number (*n*)		259	164	356	8	9
Mean depth (m)	Mean	0.123	0.265	0.375	0.610	0.403
	Minimum	0.01	0.03	0.03	0.18	0.06
	Maximum	0.83	0.76	1.63	1.89	1.15
Percent bedrock	Mean	10.0	9.0	12.9	6.3	8.8
	Minimum	0	0	0	0	0
	Maximum	100	95	95	50	50
Percent boulder	Mean	16.4	18.0	13.2	10.9	4.0
	Minimum	0	0	0	0	0
	Maximum	80	90	80	55	20
Percent silt–clay	Mean	1.0	1.5	5.0	9.6	8.9
	Minimum	0	0	0	1	0
	Maximum	45	90	70	25	40
Percent vegetation	Mean	13.2	5.4	8.7	3.8	6.9
	Minimum	0	0	0	0	0
	Maximum	99	70	95	10	30
Velocity (coded)	Mean	3	2	1	1	1
	Minimum	3	2	1	1	1
	Maximum	3	2	1	1	1
Wood (no.·m^{-2})	Mean	0.01460	0.05080	0.01415	0.00415	0.00422
	Minimum	0.0000	0.0000	0.0000	0.0000	0.0000
	Maximum	0.3682	3.8760	0.6667	0.0139	0.0180
Ozark Highlands						
Number (*n*)		105	113	92	10	46
Mean depth (m)	Mean	0.254	0.475	0.518	0.636	0.382
	Minimum	0.04	0.12	0.13	0.15	0.10
	Maximum	0.71	1.42	1.97	1.17	1.09
Percent bedrock	Mean	8.6	9.6	6.8	6.0	1.8
	Minimum	0	0	0	0	0
	Maximum	100	90	95	10	35

© 2007 NRC Canada

Table 3 (*concluded*).

Variable		Riffle	Run	Midchannel pool	Lateral pool	Backwater
Percent boulder	Mean	0.8	1.5	1.3	3.7	1.0
	Minimum	0	0	0	0	0
	Maximum	15	30	18	20	25
Percent silt–clay	Mean	2.5	3.5	6.0	2.6	30.3
	Minimum	0	0	0	0	0
	Maximum	85	80	70	10	100
Percent vegetation	Mean	5.1	3.0	3.5	2.4	12.7
	Minimum	0	0	0	0	0
	Maximum	90	80	80	15	95
Velocity (coded)	Mean	3	2	1	1	1
	Minimum	3	2	1	1	1
	Maximum	3	2	1	1	1
Wood (no.·m^{-2})	Mean	0.00764	0.01829	0.01490	0.00288	0.02148
	Minimum	0.0000	0.0000	0.0000	0.0000	0.0000
	Maximum	0.2648	0.4348	0.2302	0.0113	0.2412

Note: All variables were included in regression tree analyses.

Table 4. Pearson correlations (r) between habitat variables of 1800 channel units in 128 stream reaches in the Boston Mountains, Ouachita Mountains, and Ozark Highlands ecoregions in eastern Oklahoma.

Variable	Percent vegetation	Percent bedrock	Percent boulder	Percent silt–clay	Velocity	Wood
Mean depth	–0.053	–0.033	–0.005	0.146	–0.398	–0.046
Percent vegetation		–0.096	0.045	0.060	0.047	–0.025
Percent bedrock			–0.144	–0.135	–0.001	–0.018
Percent boulder				–0.155	0.068	0.003
Percent silt–clay					–0.223	0.008
Velocity						–0.039

smallmouth bass. Variables included in the RTA were the same as those for age-0 smallmouth bass. The regression tree with the minimum cross-validated relative error, 0.650, had 19 splitting nodes and 20 terminal nodes (Fig. 4). Shallow channel units, regardless of stream size, had the lowest densities, followed by deeper channel units in small streams. However, high densities were observed in some shallow channel units of large streams when bedrock was present. Densities in deeper channel units of small streams were higher, but only in deep channel units of wide, shallow-channel reaches or when substrates were pebble or larger. The highest densities of age-1 and older smallmouth bass were found in deep channel units of large streams that had well-developed riffle–pool morphology (<90% pool). Densities were higher in channel units intermediate in depth (0.33–0.42m) when substrates were large and channel slope was greater than 0.2%. However, the highest densities occurred in the deepest channel units where bedrock was present. When bedrock was not present, densities were higher when vegetation coverage was less than 5% and substrate composition consisted of few fines (Fig. 4).

Discussion

Densities of smallmouth bass in eastern Oklahoma streams were related to watershed morphology, reach morphology, and channel-unit characteristics in complex ways. There was no overriding influence of one spatial scale on smallmouth bass density. The influence of each scale was dependent on specific conditions at other scales, likely reflecting the linkage among scales (Frissell et al. 1986). These complex relationships were elucidated by use of regression tree analysis and would have been difficult to establish using linear modeling techniques; others have successfully used regression tree analysis to elucidate complex relationships between smallmouth bass nest densities and lake habitats (Rejwan et al. 1999). The models can also be used to predict smallmouth bass densities at a site from measured variables and guide management expectations that change channel morphology or channel-unit characteristics.

Interestingly, ecoregion did not have a direct influence on smallmouth bass density. We found that densities in streams of northeastern Oklahoma ecoregions were higher than those in the southeast, as shown in other studies (Stark and Zale 1991; Balkenbush and Fisher 2001). However, regional differences in smallmouth bass densities were more related to variation in reach morphology in larger streams with high densities than they were to ecoregions. Variation in reach morphology of large streams reflected differences in lithology, sediment regimes, land use, and hydrology among ecoregions (Woods et al. 2005; Splinter 2006), but also incorporated spatial variability among watersheds and local processes within ecoregions. Changes in land use, hydrology, and temperature that are reflective of large-scale processes have previously been shown to affect smallmouth bass populations (Peterson and Kwak 1999). Thus, although smallmouth bass abundance does differ among ecoregions, it is more closely linked with reach morphology that reflects

© 2007 NRC Canada

Table 5. Smallmouth bass densities (no. ha^{-1}) per age group per channel-unit type in streams of the Boston Mountains, Ouachita Mountains, and Ozark Highlands ecoregions in eastern Oklahoma.

		Age-0		Age-1 and older	
Channel-unit type	*n*	Mean	Range	Mean	Range
Boston Mountains					
Abandoned channel	4	278.5	0–1114.1	0.0	0–0
Backwater	31	78.7	0–1228.7	17.2	0–434
Debris pool	2	0.0	0–0	0.0	0–0
Eddy pool	2	0.0	0–0	239.6	0–479
Fall	3	0.0	0–0	0.0	0–0
Lateral pool	16	62.0	0–341	98.3	0–492
Midchannel pool	160	58.6	0–2071	143.2	0–5851
Plunge pool	3	0.0	0–0	0.0	0–0
Riffle	191	5.4	0–50	0.3	0–42
Run	183	122.3	0–9406	254.8	0–8914
Sheet	5	0.0	0–0	0.0	0–0
Ouachita Mountains					
Abandoned channel	1	0.0	0–0	0.0	0–0
Backwater	9	0.0	0–0	0.0	0–0
Eddy pool	3	0.0	0–0	0.0	0–0
Fall	4	0.0	0–0	0.0	0–0
Lateral pool	8	0.0	0–0	0.0	0–0
Midchannel pool	356	14.8	0–2096	8.0	0–775
Riffle	259	3.0	0–335	1.1	0–277
Run	164	36.1	0–2331	12.3	0–629
Ozark Highlands					
Abandoned channel	1	0.0	0–0	0.0	0–0
Backwater	46	68.7	0–1216	19.9	0–370
Debris pool	1	0.0	0–0	0.0	0–0
Eddy pool	11	241.8	0–1377	654.5	0–4 642
Fall	2	0.0	0–0	208.6	0–417
Lateral pool	10	78.2	0–463	156.2	0–458
Midchannel pool	98	16.1	0–161	146.4	0–2749
Plunge pool	2	0.0	0–0	0.0	0–0
Riffle	110	49.9	0–4873	42.8	0–4179
Run	113	40.5	0–785	160.7	0–3793
Sheet	2	0.0	0–0	0.0	0–0

ecoregion characteristics rather than being directly related to ecoregions themselves.

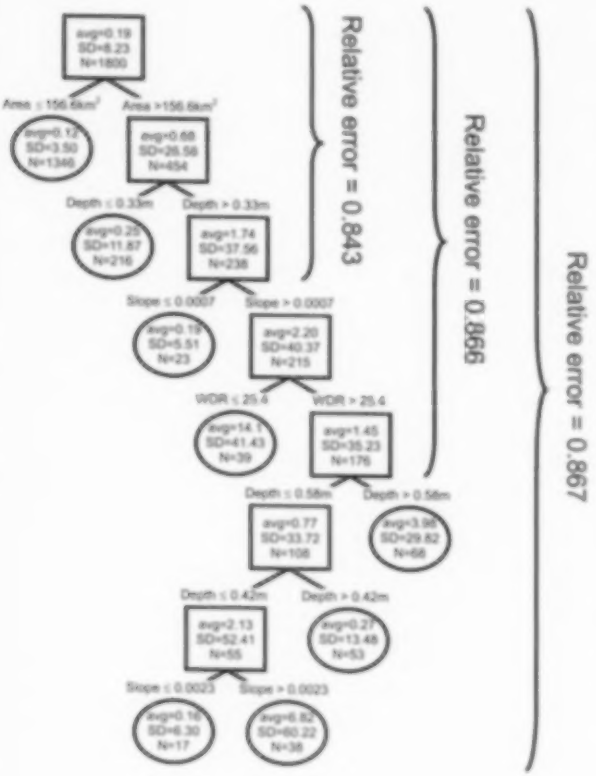
Larger streams had deeper channel units and higher densities of smallmouth bass. This likely reflected the distribution of spawning adults and spawning habitat. Small streams have stochastic stream flows and habitat that limit smallmouth bass reproduction (Orth and Newcomb 2002), and adults typically spawn in large, deep pools (Reynolds and O'Bara 1991) that are more common in larger streams. Densities were also higher in streams with wide, shallow channels. Large streams in the Ozark Highlands typically exhibited these characteristics and generally have an abundance of spawning gravels. They also have very high densities of nesting smallmouth bass (Dauwalter and Fisher 2007*b*). Floods in northeastern Oklahoma streams are also of less magnitude but longer duration than floods in southeastern streams (Dauwalter 2006), and these streamflow characteristics have been related to increased smallmouth bass abundance (Peterson and Kwak 1999).

Smallmouth bass densities did not decrease in the largest streams we sampled. This reflected our sampling design. Had we sampled larger streams (e.g., > fourth order), we likely would have observed decreased densities or absence of smallmouth bass (Stark and Zale 1991; Rabeni and Jacobson 1993); in Baron Fork Creek, smallmouth bass densities were higher in a fourth-order reach when compared with a fifth-order reach (Dauwalter 2006). Rabeni and Jacobson (1993) suggested that smallmouth bass densities in Ozark streams decrease with increased valley width owing to finer bedload materials and less bluff pools, which are important smallmouth bass habitats. Increased stream temperatures downstream may also limit smallmouth populations (Zweifel et al. 1999). Thus, intermediate-sized streams and rivers in eastern Oklahoma likely provide the streamflow and habitat conditions most suitable for smallmouth bass populations.

Channel morphology was primarily important in explaining densities of smallmouth bass in large streams and deep channel units. Width–depth ratio typically reflects channel

© 2007 NRC Canada

Fig. 3. Regression tree analysis of effects of ecoregion, watershed and reach morphology and habitat, and channel-unit habitat on age-0 smallmouth bass (*Micropterus dolomieu*) densities (no./ha^{-1}) in 1800 channel units from 128 stream reaches in three eastern Oklahoma ecoregions. Mean densities per node (avg) are given, as are standard deviation (SD) and sample size (*N*). Terminal nodes are oval. Tenfold cross-validated relative error was 0.843 in the original analysis and 0.866 and 0.867 in exploratory analyses.



stability and percent silt–clay in the channel perimeter and banks (both negatively related) and is positively related to discharge and sediment loads; sinuosity also reflects silt–clay content of banks and typically varies inversely with width–depth ratio (Knighton 1998). In the Ozark Highlands, accelerated erosion due to past clearing of forests and current overgrazing has likely increased runoff and bedload dynamics, resulting in wider, gravel-dominated channels (Rabeni and Jacobson 1993; Splinter 2006). Increased channel width also results in less canopy cover per stream width. This may increase production of stream invertebrates beneficial to fishes, such as juvenile smallmouth bass (Livingstone and Rabeni 1991), as long as the increased sediment load is primarily gravel and not fines (Waters 1995). We also found that reaches with more riffle–pool structure (less percent pool overall) in larger streams had higher smallmouth bass densities. In Missouri streams, smallmouth bass densities were higher in streams with less pool habitat (Sowa and Rabeni 1995). Crayfish production has been reported to be higher in riffles than pools in larger streams (Roell and Orth 1992). Thus, as riffle–pool ratios decrease there may be less crayfish production per reach, leading to reduced prey availability to smallmouth bass.

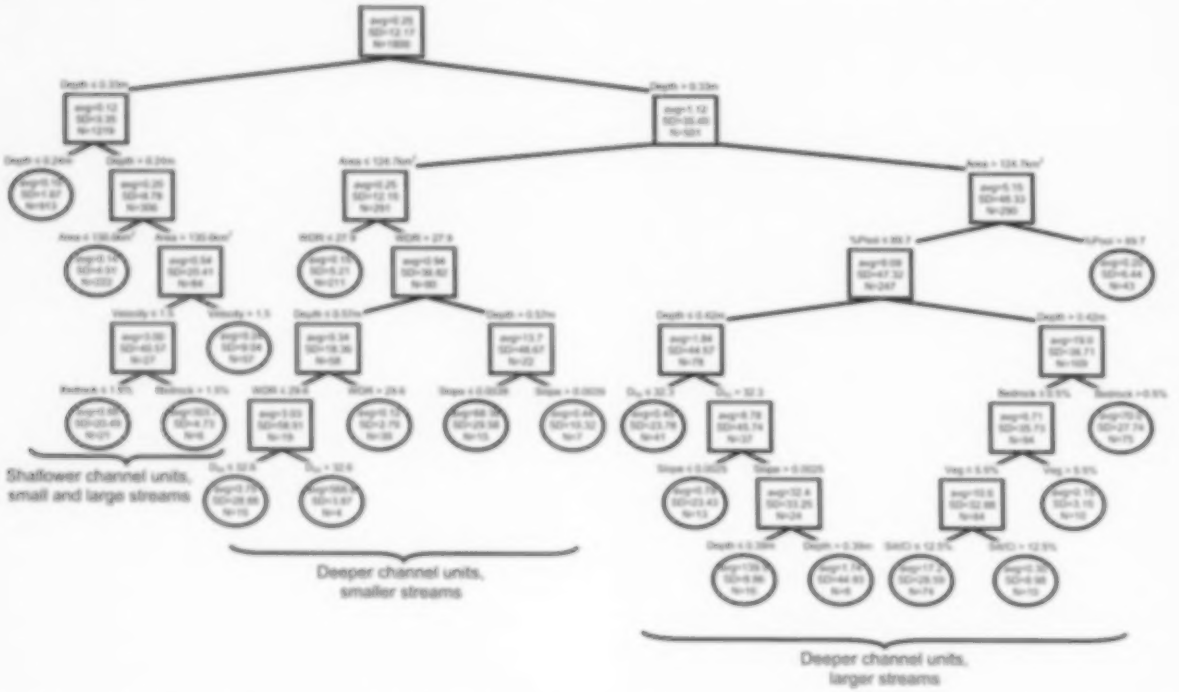
Except for water depth, channel-unit habitat was not related to age-0 densities of smallmouth bass, and habitat was mostly related to densities of age-1 and older bass in deeper channel units of larger streams. Densities of age-0 and age-1 and older smallmouth bass were higher in deeper channel units and were typically very low or zero in shallow ones regardless of stream size. Predators like smallmouth bass typically occupy deeper habitats of streams (Schlosser 1987), and they do not use depths less than about 0.25 m regardless of size (Orth and Newcomb 2002; Fore et al. 2007). Avoidance of shallow water may be a behavioral adaptation to circumvent predation by terrestrial vertebrates. Although deep channel units were present in both small and large streams, deeper channel units were more prevalent in larger streams. Consequently, densities were higher overall in the larger streams we sampled. Age-0 smallmouth bass are typically generalized in habitat use (Sabo and Orth 1994). Swimming ability of age-0 individuals may also prohibit them from moving into desired channel units, especially between pools separated by riffles. Consequently, densities of age-0 smallmouth bass in channel units may reflect the quality of spawning habitat (Cleary 1956), number of successful nests and nest production (Reynolds and O'Bara 1991; Lukas and Orth 1995), and (or) survival dynamics of fry and juveniles through their first summer (Knotek and Orth 1998).

Densities of age-1 and older smallmouth bass were only related to channel-unit habitat (other than depth) in a few instances. Densities were higher in shallow channel units of larger streams when bedrock was present. In deep channel units of large streams, densities were also higher when bedrock was present. Local bed scour occurs along bedrock outcroppings and can create deep areas in even shallow channel units. We also observed no density relationship with boulders, as was reported in Jacks Fork River, Missouri (McClendon and Rabeni 1987). Boulders have been associated with bluff pools, which are important habitats for smallmouth bass (Rabeni and Jacobson 1993). We also found that densities were lower when more vegetation was present. Use of vegetation can be detrimental to age-0 smallmouth bass when compared with other habitats (Olson et al. 2003); however, vegetation had no effect on densities in other studies (McClendon and Rabeni 1987). Channel units with a fraction of silt–clay substrate also had low smallmouth bass densities. Fine sediments represent unsuitable habitat for the benthic macroinvertebrates and crayfishes consumed by smallmouth bass in these streams (Waters 1995; Pflieger 1996; Fenner et al. 2004), and smallmouth bass have been shown to be absent more often when reaches had more fine substrates (Lyons 1991).

The effects of spatial scale need to be considered in the management and conservation of smallmouth bass. Ecoregions represent the largest scale we studied. The differences in abundance of sport fishes like smallmouth bass among ecoregions led to identification of ecoregional-based fishery management regions in eastern Oklahoma (northeast, east-central, and southeast; Fisher et al. 2004). Of the components used to delineate ecoregions, land use is the only one that can be actively managed. Management of land use

© 2007 NRC Canada

Fig. 4. Regression tree analysis of effects of ecoregion, watershed and reach morphology and habitat, and channel-unit habitat on age-1 and older smallmouth bass (*Micropterus dolomieu*) densities (no.·ha^{-1}) in 1800 channel units from 128 stream reaches in three eastern Oklahoma ecoregions. Mean densities per node (avg) are given, as are standard deviation (SD) and sample size (*N*). Terminal nodes are oval. Tenfold cross-validated relative error was 0.650. Broad descriptions of channel units related to primary splits are given.


Shallower channel units, small and large streams
Deeper channel units, smaller streams
Deeper channel units, larger streams

on a large scale (ecoregions or watersheds) is likely to be slow and problematic. However, long-term changes in land use can be beneficial to smallmouth bass populations. Conversion of land from forest or prairie to agriculture can increase flood magnitude and reduce flood pulse duration. Floods in north temperate regions occur in spring when smallmouth bass spawn, and increased flood magnitude during spawning can adversely affect population size (Peterson and Kwak 1999). The contributing watershed and subsequent size of a stream also cannot be actively managed. Management at the ecoregion and watershed scales is more likely to involve strategies tailored for specific regional or stream conditions. For example, fishing regulations may be tailored for streams in each region, and the expected catch rates of anglers and responses of fish populations to habitat management should be adjusted to regions and stream size.

Stream habitat management at the reach and channel-unit spatial scales is more logistically and fiscally feasible and has a long history (White 1996). Reach morphology is an indicator of stream condition, and stream reaches are often classified to guide instream habitat management (Rosgen 1994; Kondolf et al. 2003). Changing reach morphology is often a goal of habitat management, especially alteration of riffle–pool ratios and channel width–depth ratios (Roni et al. 2005). We found that smallmouth bass were more abundant in streams with riffle–pool morphology, and they have been shown to respond negatively to anthropogenic activities that increase channel widths, alter riffle–pool morphology, and pool depth (Brown et al. 1998). Instream structures are often used to narrow and deepen stream channels and create deep pools used by sport fishes (Binns 1994). Restoration of riparian vegetation can also promote bank stability and change channel morphology. It can also moderate water temperatures in summer (Poole and Berman 2001) and benefit growth, and possibly population size, of smallmouth bass (Whitledge et al. 2006). Addition of wood, boulders, and other habitat within channel units is another habitat management practice. Smallmouth bass are typically associated with specific microhabitats, including cover. However, our data suggest that substrate and vegetation characteristics become important to smallmouth bass only after channel-unit size, stream size, and channel morphology result in locally favorable habitat conditions; similar interactions between spatial scales were suggested for Missouri streams (McClendon and Rabeni 1987; Rabeni and Sowa 1996). Thus, microhabitats likely influence the spatial distribution of smallmouth bass in a channel unit but do not strongly influence production or population dynamics like larger scale factors do (Rabeni and Sowa 1996), and although addition of instream cover can sometimes be useful to effect change in fish populations and communities (House 1996), it should not replace maintenance or re-establishment of the large-scale processes of water, sediment, and wood delivery that create reach characteristics favorable to smallmouth bass populations (Poole et al. 1997).

The need to understand how physical characteristics of streams influence stream fishes at multiple spatial scales is not a new concept (Frissell et al. 1986; Wiens 2002) but is

© 2007 NRC Canada

essential for effective stream management. We observed strong stream-size effects, but geomorphic influences on smallmouth bass were complex and different between small and large streams. Rabeni and Sowa (1996) stressed the importance of realizing the nestedness of spatial scales and understanding how factors acting on different spatial scales are important to the conservation of stream fishes. Geomorphic processes are naturally hierarchical, and understanding spatial variability associated with those processes is important to understanding how they structure fish habitat in streams (Montgomery 1999). Most often, large-scale processes influence stream morphology and function at smaller spatial scales, but local impacts do not always influence large-scale processes. There are exceptions, as upstream and downstream channel adjustments may occur from a localized disturbance. Consequently, many stream restoration and habitat enhancement projects have a geomorphic basis that includes different spatial scales (Rosgen 1996; Federal Interagency Stream Restoration Working Group 1998; Montgomery and Bolton 2003). Understanding how smallmouth bass populations relate to large-scale factors, channel morphology, and channel-unit habitat, as shown here, is essential in realizing how certain conservation and management practices might influence this popular sport fish (Fisher et al. 2002) and indicator species (Hlass et al. 1998) of these upland regions in the central United States.

Acknowledgments

We thank V. Horncastle, S. Davis, A. Krystyniak, B. Marston, R. Ary, K. Winters, M. Murray, and J. Morel for field assistance. J. Bidwell, A. Echelle, and two anonymous reviewers commented on manuscript drafts. Project funding was provided by a Federal Aid in Sport Fish Restoration Act grant under Project F-55-R of the Oklahoma Department of Wildlife Conservation and the Oklahoma Cooperative Fish and Wildlife Research Unit. The Oklahoma Cooperative Fish and Wildlife Research Unit is jointly sponsored by the US Geological Survey, Oklahoma State University, the Oklahoma Department of Wildlife Conservation, the Wildlife Management Institute, and the US Fish and Wildlife Service. D. Dauwalter was supported by a fellowship for Water, Energy, and the Environment from the Environmental Institute, Oklahoma State University.

References

Bain, M.B. 1999. Substrate. *In* Aquatic habitat assessment: common methods. *Edited by* M.B. Bain and N.J. Stevenson. American Fisheries Society, Bethesda, Maryland. pp. 95–103.

Balkenbush, P.E., and Fisher, W.L. 2001. Population characteristics and management of black bass in eastern Oklahoma streams. Proceedings of the Southeastern Association of Fish and Wildlife Agencies, **53**: 130–143.

Baxter, C.V., and Hauer, F.R. 2000. Geomorphology, hyporheic exchange, and selection of spawning habitat by bull trout (*Salvelinus confluentus*). Can. J. Fish. Aquat. Sci. **57**: 1470–1481.

Benda, L., Poff, N.L., Miller, D., Dunne, T., Reeves, G., Pess, G., and Pollock, M. 2004. The network dynamics hypothesis: how channel networks structure riverine habitat. Bioscience, **54**: 413–427.

Beschta, R.L., and Platts, W.S. 1986. Morphological features of small streams: significance and function. Water Resour. Bull. **22**: 370–379.

Binns, N.A. 1994. Long-term responses of trout and macrohabitats to habitat management in a Wyoming headwater stream. N. Am. J. Fish. Manag. **14**: 87–98.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. 1984. Classification and regression trees. Wadsworth International Group, Belmont, California.

Brown, A.V., Lyttle, M.M., and Brown, K.B. 1998. Impacts of gravel mining on gravel bed streams. Trans. Am. Fish. Soc. **127**: 979–994.

Champoux, O., Biron, P.M., and Roy, A.G. 2003. The long-term effectiveness of fish habitat restoration practices: Lawrence Creek, Wisconsin. Ann. Assoc. Am. Geographers, **93**: 42–54.

Cleary, R.E. 1956. Observations on factors affecting smallmouth bass production in Iowa. J. Wildl. Manag. **20**: 353–359.

Dauwalter, D.C. 2006. Relationships among geomorphology, habitat, and fishes in eastern Oklahoma streams: implications for stream restoration. Ph.D. thesis, Oklahoma State University, Stillwater, Oklahoma.

Dauwalter, D.C., and Fisher, W.L. 2007*a*. Electrofishing capture probability of smallmouth bass in streams. N. Am. J. Fish. Manag. **27**: 162–171.

Dauwalter, D.C., and Fisher, W.L. 2007*b*. Spawning chronology, nest site selection, and nest success of smallmouth bass during benign streamflow conditions. Am. Midl. Nat. **158**: 60–78.

Dauwalter, D.C., Fisher, W.L., and Belt, K.C. 2006. Mapping stream habitats with a global positioning system: accuracy, precision, and comparison with traditional methods. Environ. Manag. **37**: 271–280.

De'ath, G., and Fabricius, K.E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology, **81**: 3178–3192.

Dolloff, A., Kershner, J., and Thurow, R. 1996. Underwater observation. *In* Fisheries techniques. *Edited by* B.R. Murphy and D.W. Willis. American Fisheries Society, Bethesda, Maryland. pp. 533–554.

Fenner, D.B., Walsh, M.G., and Winkelman, D.L. 2004. Diet overlap of introduced rainbow trout and three native fishes in an Ozark stream. Am. Fish. Soc. Symp. **44**: 475–482.

Fisher, W.L., Schreiner, D.F., Martin, C.D., Negash, Y.A., and Kessler, E. 2002. Recreational fishing and socioeconomic characteristics of eastern Oklahoma stream anglers. Proc. Okla. Acad. Sci. **82**: 79–87.

Fisher, W.L., Tejan, E.C., and Balkenbush, P.E. 2004. Regionalization of stream fisheries management using Geographic Information Systems. *In* Proceedings of the 2nd International Symposium on GIS/Spatial Analyses in Fishery and Aquatic Sciences. *Edited by* T. Nishida, P.J. Kailola, and C.E. Hollingworth. Fishery–Aquatic GIS Research Group, Kawagoe, Saitama, Japan. pp. 465–476.

Federal Interagency Stream Restoration Working Group. 1998. Stream corridor restoration: principles, processes, and practices. The Federal Interagency Stream Restoration Working Group (FISRWG) (15 federal agencies of the US gov't). GPO Item No. 0120; SuDocs No. 57.6/2:EN3/PT.653. ISBN-0-934213-59-3.

Fore, J.D., Dauwalter, D.C., and Fisher, W.L. 2007. Microhabitat use by smallmouth bass in an Ozark stream. J. Freshw. Ecol. **22**(2): 189–199.

Frissell, C.A., Liss, W.J., Warren, C.E., and Hurley, M.D. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environ. Manag. **10**: 199–214.

Frothingham, K.M., Rhoads, B.L., and Herricks, E.E. 2002. A multiscale conceptual framework for integrated ecogeomorpho-

© 2007 NRC Canada

logical research to support stream naturalization in the agricultural Midwest. Environ. Manag. **29**: 16–33.

Gido, K.B., Falke, J.A., Oakes, R.M., and Hase, K.J. 2006. Fish-habitat relations across spatial scales in prairie streams. *In* Landscape influences on stream habitats and biological assemblages. *Edited by* R.M. Hughes, L. Wang, and P.W. Seelbach. Am. Fish. Soc. Symp. **48**: 265–285.

Hawkins, C.P., Kershner, J.L., Bisson, P.A., Bryant, M.D., Decker, L.M., Gregory, S.V., McCullough, D.A., Overton, C.K., Reeves, G.H., Steedman, R.J., and Young, M.K. 1993. A hierarchical approach to classifying stream habitat features. Fisheries, **18**(6): 3–12.

Hlass, L.J., Fisher, W.L., and Turton, D.J. 1998. Use of the index of biotic integrity to assess water quality in forested streams in the Ouachita Mountains ecoregion, Arkansas. J. Freshw. Ecol. **13**: 181–192.

House, R. 1996. An evaluation of stream restoration structures in a coastal Oregon stream 1981–1993. N. Am. J. Fish. Manag. **16**: 272–281.

Huryn, A.D., and Wallace, J.B. 1987. Local geomorphology as a determinant of macrofaunal production in a mountain stream. Ecology, **68**: 1932–1942.

Jackson, D.A. 1993. Stopping rules in PCA: a comparison of heuristical and statistical approaches. Ecology, **74**: 2205–2214.

Knighton, D. 1998. Fluvial forms and processes: a new perspective. Oxford University Press, Inc., New York.

Knotek, W.L., and Orth, D.J. 1998. Survival for specific life intervals of smallmouth bass, *Micropterus dolomieu*, during parental care. Environ. Biol. Fishes, **51**: 285–296.

Kondolf, G.M., Montgomery, D.R., Piéjay, H., and Schmitt, L. 2003. Geomorphic classification of rivers and streams. *In* Tools in geomorphology. *Edited by* G.M. Kondolf and H. Piéjay. John Wiley & Sons Ltd., West Sussex, England. pp. 171–204.

Livingstone, A.C., and Rabeni, C.F. 1991. Food–habitat relations of underyearling smallmouth bass in an Ozark stream. *In* The First International Smallmouth Bass Symposium. *Edited by* D.C. Jackson. Mississippi Agricultural and Forestry Experiment Station, Mississippi State University, Mississippi State, Miss. pp. 76–83.

Lukas, J.A., and Orth, D.J. 1995. Factors affecting nesting success of smallmouth bass in a regulated Virginia stream. Trans. Am. Fish. Soc. **124**: 726–735.

Lyons, J. 1991. Predicting smallmouth bass presence/absence and abundance in Wisconsin streams using physical habitat characteristics. *In* The First International Smallmouth Bass Symposium. *Edited by* D.C. Jackson. Mississippi Agricultural and Forestry Experiment Station, Mississippi State University, Mississippi State, Miss. pp. 96–103.

McClendon, D.D., and Rabeni, C.F. 1987. Physical and biological variables useful for predicting population characteristics of smallmouth bass and rock bass in an Ozark stream. N. Am. J. Fish. Manag. **7**: 46–56.

McGarigal, K., Cushman, S., and Stafford, S. 2000. Multivariate statistics for wildlife and ecology research. Springer-Verlag, New York.

Miranda, L.E., and Dolan, C.R. 2003. Test of a power transfer model for standardizing electrofishing. Trans. Am. Fish. Soc. **132**: 1179–1185.

Montgomery, D.R. 1999. Process domains and the river continuum. J. Am. Water Resour. Assoc. **36**: 397–410.

Montgomery, D.R., and Bolton, S.M. 2003. Hydrogeomorphic variability and river restoration. *In* Strategies for restoring river ecosystems: sources of variability and uncertainty in natural and managed systems. *Edited by* R.C. Wissmar and P.A. Bisson. American Fisheries Society, Bethesda, Maryland. pp. 39–80.

Morisawa, M. 1968. Streams: their dynamics and morphology. McGraw-Hill Book Company, New York.

Newson, M.D., and Newson, C.L. 2000. Geomorphology, ecology and river channel habitat: mesoscale approaches to basin-scale challenges. Prog. Phys. Geog. **24**: 195–217.

Olson, M.H., Young, B.P., and Blinkoff, K.D. 2003. Mechanisms underlying habitat use of juvenile largemouth bass and smallmouth bass. Trans. Am. Fish. Soc. **132**: 398–405.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Ann. Assoc. Am. Geographers, **77**: 118–125.

Orth, D.J., and Newcomb, T.J. 2002. Certainties and uncertainties in defining essential habitats for riverine smallmouth bass. *In* Black bass: ecology, conservation, and management. *Edited by* D.P. Philipp and M.S. Ridgway. Am. Fish. Soc. Symp. **31**: 251–264.

Peterson, J.T., and Kwak, T.J. 1999. Modeling the effects of land use and climate change on riverine smallmouth bass. Ecol. Appl. **9**: 1391–1404.

Pflieger, W.L. 1996. The crayfishes of Missouri. Missouri Department of Conservation, Jefferson City, Missouri.

Poole, G.C., and Berman, C.H. 2001. An ecological perspective on in-stream temperature: natural heat dynamics and mechanisms of human-caused thermal degradation. Environ. Manag. **27**: 787–802.

Poole, G.C., Frissell, C.A., and Ralph, S.C. 1997. In-stream habitat unit classification: inadequacies for monitoring and some consequences for management. J. Am. Water Resour. Assoc. **33**: 879–896.

Rabeni, C.F., and Jacobson, R.B. 1993. Geomorphic and hydraulic influences on the abundance and distribution of stream centrarchids in Ozark USA streams. Pol. Arch. Hydrobiol. **40**: 87–99.

Rabeni, C.F., and Sowa, S.P. 1996. Integrating biological realism into habitat restoration and conservation strategies for small streams. Can. J. Fish. Aquat. Sci. **53**(Suppl. 1): 252–259.

Rejwan, C., Collings, N.C., Brunner, L.J., Shuter, B.J., and Ridgway, M.S. 1999. Tree regression analysis on the nesting habitat of smallmouth bass. Ecology, **80**: 341–348.

Reynolds, C.R., and O'Bara, C.J. 1991. Reproductive ecology and spawning habitat of smallmouth bass in two small streams of the Tennessee River system. *In* The First International Smallmouth Bass Symposium. *Edited by* D.C. Jackson. Mississippi Agricultural and Forestry Experiment Station, Mississippi State, Miss. pp. 61–65.

Rhoads, B.L., Schwartz, J.S., and Porter, S. 2003. Stream geomorphology, bank vegetation, and three-dimensional habitat hydraulics for fish in midwestern agricultural streams. Water Resour. Res. **39**: 1218.

Roell, M.J., and Orth, D.J. 1992. Production of three crayfish populations in the New River of West Virginia, USA. Hydrobiologia, **228**: 185–194.

Roni, P., Fayram, A.H., and Miller, M.A. 2005. Monitoring and evaluating instream habitat enhancement. *In* Monitoring stream and watershed restoration. *Edited by* P. Roni. American Fisheries Society, Bethesda, Maryland. pp. 209–236.

Rosgen, D.L. 1994. A classification of natural rivers. Catena, **22**: 169–199.

Rosgen, D.L. 1996. Applied river morphology. Wildland Hydrology, Pagosa Springs, Colorado.

Rutherford, D.A., Echelle, A.A., and Maughan, O.E. 1992. Drainage-wide effects of timber harvesting on the structure of stream fish assemblages in southeastern Oklahoma. Trans. Am. Fish. Soc. **121**: 716–728.

Sabo, M.J., and Orth, D.J. 1994. Temporal variation in microhabitat use by age-0 smallmouth bass in the North Anna River, Virginia. Trans. Am. Fish. Soc. **123**: 733–746.

© 2007 NRC Canada

Schlosser, I.J. 1987. A conceptual framework for fish communities in small warmwater streams. *In* Community and evolutionary ecology of North American stream fishes. *Edited by* W.J. Matthews and D.C. Heins. University of Oklahoma Press, Norman, Okla. pp. 17–24.

Sowa, S.P., and Rabeni, C.F. 1995. Regional evaluation of the relation of habitat to distribution and abundance of smallmouth bass and largemouth bass in Missouri streams. Trans. Am. Fish. Soc. **124**: 240–251.

Splinter, D.K. 2006. Spatial patterns in the fluvial system: comparisons among three eastern Oklahoma ecoregions. Ph.D. thesis, Oklahoma State University, Stillwater, Oklahoma.

Stark, W.J., and Zale, A.V. 1991. Status of smallmouth bass populations in eastern Oklahoma. Oklahoma Department of Wildlife Conservation, Oklahoma City, Federal Aide Project No. F-41-R.

Steinberg, D., and Colla, P. 1995. CART: tree-structured nonparametric data analysis. Salford Systems, San Diego, California.

Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. Trans. Am. Geophys. Union, **38**: 913–920.

US Geological Survey. 1999. National elevation dataset. 1st ed. US Geological Survey, Sioux Falls, South Dakota.

Vannote, R.L., Minshall, G.W., Cummins, K.W., and Cushing, C.E. 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. **37**: 130–137.

Walters, D.M., Leigh, D.S., Freeman, M.C., Freeman, B.J., and Pringle, C.M. 2003. Geomorphology and fish assemblages in a Piedmont river basin, U.S.A. Freshw. Biol. **48**: 1950–1970.

Waters, T.F. 1995. Sediment in streams: sources, biological effects, and control. American Fisheries Society Monograph 7, Bethesda, Maryland.

White, R.J. 1996. Growth and development of North American stream habitat management for fish. Can. J. Fish. Aquat. Sci. **53**(Suppl. 1): 342–363.

Whitledge, G.W., Rabeni, C.F., Annis, G., and Sowa, S.P. 2006. Riparian shading and groundwater enhance growth potential for smallmouth bass in Ozark streams. Ecol. Appl. **16**: 1461–1473.

Wiens, J.A. 2002. Riverine landscapes: taking landscape ecology into the water. Freshw. Biol. **47**: 501–515.

Woods, A.J., Omernik, J.M., Butler, D.R., Ford, J.G., Henley, J.E., Hoagland, B.W., Arndt, D.S., and Moran, B.C. 2005. Ecoregions of Oklahoma (color poster with map, descriptive text, summary tables, and photographs). US Geological Survey, Reston, Virginia.

Yang, C.T. 1971. Formation of riffles and pools. Water Resour. Res. **7**: 1567–1574.

Zar, J.H. 1999. Biostatistical analysis. Prentice Hall, Upper Saddle River, New Jersey.

Zweifel, R.D., Hayward, R.S., and Rabeni, C.F. 1999. Bioenergetics insight into black bass distribution shifts in Ozark border region streams. N. Am. J. Fish. Manag. **19**: 192–197.

© 2007 NRC Canada

Estimation of the parameters of fish stock dynamics from catch-at-age data and indices of abundance: can natural and fishing mortality be separated?

Sondre Aanes, Steinar Engen, Bernt-Erik Sæther, and Ronny Aanes

Abstract: Models for fluctuations in size of fish stocks must include parameters that describe expected dynamics, as well as stochastic influences. In addition, reliable population projections also require assessments about the uncertainties in estimates of vital parameters. Here we develop an age-structured model of population dynamics based on catch-at-age data and indices of abundance in which the natural and fishing mortality are separated in a Bayesian state–space model. Markov chain Monte Carlo methods are used to fit the model to the data. The model is fitted to a data set of 19 years for Northeast Arctic cod (*Gadus morhua*). By simulations of the fitted model we show that the model captures the dynamical pattern of natural mortality adequately, whereas the absolute size of natural mortality is difficult to estimate. Access to long time series of high-quality data are necessary for obtaining precise estimates of all the parameters in the model, but some parameters cannot be estimated without including some prior information. Nevertheless, our model demonstrates that temporal variability in natural mortality strongly affects perceived variability in stock sizes. Thus, using estimation procedures that neglect temporal fluctuations in natural mortality may therefore give biased estimates of fluctuations in fish stock sizes.

Résumé : Les modèles fluctuation de la taille des stocks de poissons doivent inclure des paramètres qui décrivent la dynamique attendue ainsi que les influences stochastiques. De plus, des projections fiables de la population nécessitent une évaluation des incertitudes dans les estimations des paramètres vitaux. Nous mettons au point un modèle de la dynamique de la population structuré en fonction de l'âge qui est basé sur les données de captures en fonction de l'âge et les indices d'abondance dans lequel la mortalité naturelle et la mortalité due à la pêche sont séparées dans un modèle état-espace bayésien. Des méthodes de Monte Carlo par chaîne de Markov servent à ajuster le modèle aux données. Nous ajustons le modèle à une matrice de données récoltées sur 19 années chez la morue (*Gadus morhua*) du nord-est de l'Arctique. Par simulations du modèle ajusté, nous montrons que le modèle capture adéquatement le patron dynamique de la mortalité naturelle, bien que la taille absolue de la mortalité naturelle soit difficile à estimer. Il est nécessaire d'avoir accès à de longues séries chronologiques de données de haute qualité pour obtenir des estimations précises de tous les paramètres du modèle, mais certains paramètres ne peuvent être estimés sans l'inclusion de renseignements a priori. Néanmoins, notre modèle démontre que la variabilité temporelle de la mortalité naturelle affecte fortement la variabilité perçue des tailles des stocks. Ainsi, l'utilisation de procédures qui négligent les fluctuations temporelles de la mortalité naturelle peut donner des estimations erronées des fluctuations dans les tailles des stocks de poissons.

[Traduit par la Rédaction]

Introduction

A sustainable management strategy of commercial fish stocks is dependent on understanding how harvesting may influence future stock sizes. This requires that we are able to estimate parameters such as density and causes of mortality with sufficient precision. Accordingly, we must also estimate the stochastic influences on the population dynamics, in most cases arising from stochastic variation in the environment (Lande et al. 2003). It is important to separate the environmental noise from the stochasticity that is caused by uncertainties in parameter estimates and errors in estimates of stock size (Freckleton et al. 2006). Developing population models of some credibility requires that both types of sto-

Received 19 June 2006. Accepted 12 April 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 9 August 2007. J19372

S. Aanes.[1] Institute of Marine Research, P.O. Box 1870, Nordnes, N-5817 Bergen, Norway.
S. Engen. Norwegian University of Science and Technology, Department of Mathematical Sciences, N-7491 Trondheim, Norway.
B.-E. Sæther. Norwegian University of Science and Technology, Department of Biology, N-7491 Trondheim, Norway.
R. Aanes. Norwegian Polar Institute, Polar Environmental Centre, N-9296 Tromsø, Norway.

[1]Corresponding author (e-mail: sondre.aanes@imr.no).

 doi:10.1139/F07-074 © 2007 NRC Canada

chastic influences are properly assessed when making predictions that form the basis for management decisions (Walters and Martell 2004).

Catch-at-age data probably represent the most important data source for assessment of fish stock size (see reviews in, e.g., Megrey (1989), Quinn and Deriso (1999), and Millar and Meyer (2000)). Recent analyses of catch-at-age data have been refined to consider process and measurement errors using traditional frequentist techniques (e.g., Gudmundsson 1994; Schnute and Richards 1995; Richards and Schnute 1998), as well as Bayesian methods (e.g., Schnute 1994; Punt and Hilborn 1997; Millar and Meyer 2000). Information about the precision in the data is necessary for obtaining unbiased estimates of population parameters (de Valpine and Hastings 2002; Dennis et al. 2006; and Freckleton et al. 2006, and references therein). Furthermore, whatever the statistical approach used, separating natural fluctuations in mortality from mortality due to fishing is crucial for obtaining reliable estimates of stock size and detecting changes in harvest rates (Walters and Martell 2004).

It is commonly assumed that natural mortality of adults is known and constant (e.g., Hilborn and Walters 1992; Quinn and Deriso 1999; International Council for the Exploration of the Sea (ICES) 2005). Even if wrong values of natural mortality are chosen and if fluctuations in total mortality are dominated by fishing mortality, estimates of stock size, although potentially seriously biased, may still capture the temporal dynamics in the stock. Methods for estimating natural mortality based on capture–recapture analyses of tagged individuals (Seber 1982; Quinn and Deriso 1999) are often of limited use for marine fish populations because of high costs involved in the necessary sampling (e.g., tagging of fish). Most of the applied methods are based on samples of stomach contents and estimation of additional mortality by predation of other species in a multispecies context (ICES 2005) or, as for the Northeast Arctic cod, by cannibalism (ICES 2004).

The stochastic processes in the population dynamics along with the uncertainty in the available data for assessing fish stocks emphasises the need of probabilistic approaches for quantifying the risks concerning status of stock and management decisions. One model class capable of addressing numerous sources of variability is the state–space framework, i.e., the available data (the estimates of catch at age and indices) are noisy observations from a hidden dynamical system (the population). For instance, Gudmundsson (1994) models the fishing mortality as a stochastic process where natural mortality is assumed known and constant, whereas Lewy and Nielsen (2003) assume that stock survival (the sum of natural and fishing mortality) for a given mortality rate is a stochastic process. The variability in fishing mortality is thus confounded with the variability in natural mortality (Lewy and Nielsen 2003). In the models of Schnute and Richards (1995) and Millar and Meyer (2000), natural mortality is considered unknown, but the stochastic variations in natural mortality are not separated from the variations in fishing mortality.

In this paper we present a model in a state–space framework. The model is age-structured, and both natural and fishing mortality are considered as stochastic processes with unknown parameters. The model takes into account that the input data are uncertain. We use a Bayesian approach and Markov chain Monte Carlo methods to fit the model to the data (see Clark 2005). Starting with a complex model, we show how model complexity must be reduced (e.g., Ludwig and Walters 1989; Richards and Schnute 1998) to obtain reliable estimates of population sizes, as well as population dynamical parameters. The model is fitted to data for Northeast Arctic cod (*Gadus morhua*) in the Barents Sea, and the fit is evaluated by simulations and compared with existing estimates for this stock.

Model

Fish stock dynamics are typically nonlinear and high-dimensional because of age structure and age-specific dynamical processes. These factors complicate traditional use of maximum likelihood for fitting state–space models to data. Linear models can be fitted using the Kalman filter and maximum likelihood when the errors are normally distributed (Schnute and Richards 1995). Methods for handling nonlinearity do exist, and Gudmundsson (1994) applied the extended Kalman filter to approximate the nonlinear likelihood function. Similarly, de Valpine and Hilborn (2005) and Dennis et al. (2006) fit nonlinear state–space models by a maximum likelihood approach, but without taking into account age structure. The Bayesian approach represents an alternative to the frequentist approach and has been applied increasingly within fisheries in the last decades (e.g., Patterson 1999; Millar and Meyer 2000; Lewy and Nielsen 2003). The Bayesian approach relies on evaluation of the posterior probability distribution $P(\boldsymbol{\Theta}|\mathbf{Y})$ of the set of parameters $\boldsymbol{\Theta}$ given data $\mathbf{Y}$, prior distribution $P(\boldsymbol{\Theta})$ of the parameters, and a likelihood function $P(\mathbf{Y}|\boldsymbol{\Theta})$. By applying Bayes theorem, the posterior distribution connects the prior distribution and the likelihood function through $P(\boldsymbol{\Theta}|\mathbf{Y}) \propto P(\mathbf{Y}|\boldsymbol{\Theta})P(\boldsymbol{\Theta})$. For nonlinear and high-dimensional models, analytical evaluation of the posterior has been difficult. However, the last decade's development in computer power and numerical algorithms for sampling from the posterior, so-called Markov chain Monte Carlo (MCMC) methods (e.g., Gilks and Wild 1992; Tierney 1998; Gelman et al. 2004), has made the framework more tractable, e.g., nonlinearity does not offer any additional complexity as the distributions are evaluated numerically, and fitting complex and nonlinear models has become easier owing to the generality of the methods.

A stochastic age-structured model is developed that considers mortality as a stochastic process, decomposed into mortality caused by reported catches and natural mortality. It should be emphasised that the natural mortality must be interpreted with some care, because it also includes any unobserved fishery-induced mortality, e.g., possible discards and unreported catches (ICES (2004) and Zeller and Pauly (2005) and references therein). Random components describing the variability in each of the components are explicitly formulated.

Our approach is based on obtaining information of the hidden population dynamical process, described by the state model, from an observation model (for a similar approach, see Clark (2005), Clark and Bjørnstad (2004), and Sæther et al. (2007)). The observations are connected to the population

© 2007 NRC Canada

processes through models including error terms, allowing them to be uncertain. It is important to notice that this error in principle can not be interpreted as sampling error alone as it incorporates any deviation caused by misspecification of the functional relationship related to the true dynamical processes (see, for example, Lewy and Nielsen (2003) for a discussion).

The population dynamical model

The population size at age a in the beginning of year y, $N_{a,y}$, is connected to the population size of the same cohort the next year through

$$(1) \qquad N_{a+1,y+1} = N_{a,y}\exp(-Z_{a,y}) \qquad \text{for } a_{\min} \le a \le A, 1 \le y \le Y$$

where $Z_{a,y}$ is the mortality rate for age a in year y such that $\exp(-Z_{a,y})$ is the proportion of the cohort surviving from one year to the next. The minimum and maximum ages considered are $a_{\min}$ and A, respectively, and the years are indexed from 1 to Y, i.e., from the first to the last year considered. The mortality at age a in year y is the sum of mortality caused by fishing ($F_{a,y}$) and by natural causes ($M_{a,y}$), and thus $Z_{a,y} = F_{a,y} + M_{a,y}$.

We assume that the natural mortality at age a in year y is log-normally distributed with mean m_a. The variability in the natural mortality is decomposed into two factors: one component operating equally on all ages within year y to account for the temporal fluctuating environmental conditions influencing all ages (U_y) and the other operating independent of age and year to account for the remaining variability in natural mortality ($V_{a,y}$). More specifically,

$$(2) \qquad \log(M_{a,y}) = m_a + U_y + V_{a,y}$$

where the U_ys and $V_{a,y}$s are iid normally distributed with zero means and variances σ_U^2 and σ_M^2, respectively. This means that the covariance between the log natural mortalities at different ages is σ_U^2, and thus the correlation in log natural mortality between different ages within the year is

$$r_M = \sigma_U^2/(\sigma_U^2 + \sigma_M^2)$$

The log fishing mortality at age a in year y is modelled as a separable model with random noise (Gudmundsson 1994):

$$(3) \qquad \log(F_{a,y}) = f_a + e_y + W_{a,y}$$

where the age-specific effect (f_a) is known as the log selectivity of the fishing fleet, the year-specific effect (e_y) indicates the log effort of the fishing fleet (e.g., Quinn and Deriso 1999), and the $W_{a,y}$s are iid normally distributed with zero mean and variance σ_F^2. Notice that the e_ys are parameters in the model and not known effort data from the fishing fleet. The selectivity at age is a measure of the mean relative proportion caught by the fishing fleet at age. Above a certain age, A^*, this quantity is set constant (when the fish is fully recruited into the fishery), $f_a = f_{A^*}$ for $a \ge A^*$. We notice that the separable model may not be applicable to all types of data if, e.g., gradual changes in the structure of the fishing fleet cause systematic changes in the selectivity. The effort of the fishing fleet is not likely to show abrupt changes with time, but rather slow fluctuations caused by gradual changes in behaviour, development, and regulations of the fishing fleet. Accordingly, we model the effort as a random walk following a normal distribution, i.e.,

$$(4) \qquad e_y = e_{y-1} + T_y$$

where T_y is normally distributed with zero mean and variance σ_E^2. We fix the effort for the first year to 0 ($e_y = 0$) for identifiability (for a similar approach, see Gudmundsson 1994). If the changes in the effort employed by the fishing fleet are small from year to year, then the variance in the random walk (σ_E^2) is also small. For example, consider the special case in which there is no temporal variability in the effort, i.e., $\sigma_E^2 = 0$, then the effort is constant and fixed over the time period such that the mean log fishing mortality per age group also is constant.

The model was initialised with the recruits $N_{a_{\min},y}$ (for $y = 1, \ldots, Y$) and the initial values ($N_{a,1}$ for $a = a_{\min} + 1, \ldots, A$) that are parameters to be estimated (see also Gudmundsson 1994; Millar and Meyer 2000). A model for the recruitment could be integrated into this model and the parameters estimated simultaneously, for example, as a stock–recruitment relationship (Lewy and Nielsen 2003) or as a time series model (Schnute and Richards 1995), but such approaches are not considered in this paper.

The observations

The observations related to the population processes considered in this paper are the catch at age and the indices of abundance at age. The observed catches at age a in year y ($C_{a,y}$) are related to the population and fishing processes through

$$(5) \qquad C_{a,y} = \frac{F_{a,y}}{F_{a,y} + M_{a,y}} N_{a,y}(1 - \exp(-F_{a,y} - M_{a,y}))\lambda_{a,y}$$

where $\lambda_{a,y}$ is a stochastic variable with mean 1, so that mean catch at age equals the well-known deterministic relationship between the population and catch (e.g., Quinn and Deriso 1999). Here $\lambda_{a,y}$ represents random deviation between observed catches and true catches that is caused by sampling error and misspecification of the functional relationship (for a similar approach, see Lewy and Nielsen 2003). We assume that the catches are lognormally distributed and that the observations are independent between years and ages. More specifically, we assume $\lambda_{a,y} = \exp(\varepsilon_{a,y}^C - 0.5\sigma_C^2)$, where $\varepsilon_{a,y}^C$ is iid with zero mean and variance σ_C^2, so that $E(\lambda_{a,y}) = 1$ for any value of σ_C^2. The choice of a lognormal distribution for the catches implies that we assume a constant error coefficient of variation (ECV) for all ages given by ECV $= \sqrt{\exp(\sigma_C^2) - 1}$. An alternative approach would have been to use multinomial distributions for the proportions at age (Quinn and Deriso 1999, p. 335). However, analyses of sampling error in catch-at-age data do not support multinomial distributions (S. Aanes, unpublished data).

The second source of data considered here are the indices from the research surveys. Let δ be the proportion of the year passed until the time of the survey. Then the population size at the time of the survey is $N_{a,y}^{(\delta)} = N_{a,y}\exp(-\delta Z_{a,y})$. The estimated survey index at age a in year y ($I_{a,y}$) is assumed to

© 2007 NRC Canada

follow a lognormal distribution with mean proportional to the abundance at the time of the survey:

$$(6) \qquad I_{a,y} = q_a N^{(I)}_{a,y} \exp(\varepsilon^I_{a,y}) \quad \text{for } a^{(I)}_{\min} \le a \le A^{(I)},\ 1 \le y \le Y$$

where q_a is an age-specific proportionality constant, the "catchability" of the survey, and $\varepsilon^I_{a,y}$ is a normally distributed random variable with zero mean and variance σ^2_I, allowing for errors in the observations. The error term is assumed to be independent between ages and years. The mean of the error for the catch-at-age data was adjusted to 1 to obtain the deterministic relationship for mean catch at age, while the term of the error term for the indices is $\exp(0.5\sigma^2_I)$. For the abundance indices, such an adjustment will only affect the value of q_a and is therefore not necessary to formulate explicitly.

Data sources

In this paper we use data for Northeast Arctic cod for the time period 1985–2003. We use the annual estimates of total catch at age as reported by ICES (2004) for ages 3–15 years in addition to an age-specific survey index based on a Norwegian survey in the Barents Sea covering the feeding grounds (ages 3–9 years).

The catch at age is estimated by combining biological samples with the reported landings given by the management authorities (e.g., Aanes and Pennington 2003; Hirst et al. 2004, 2005). Because of the heterogeneous fishery for many fish stocks, possibly including a variety of fleet groups and fishing over large areas, the sampling schemes become rather complex. Accordingly, Aanes and Pennington (2003) and Hirst et al. (2004) found that the uncertainties in the estimates of the Norwegian catch at age for Northeast Arctic cod were relatively large, despite intensive sampling, with ECV typically ranging from 5% to 40%, where the least precise estimates typically are for oldest fish with the lowest catches. Because of this we choose to estimate two variance components for these data: one for the catches including ages comprising the major part of the catches (σ^2_{C1}), i.e., for ages 3–10, and one for the for catches including the ages accounting for the remaining catches (σ^2_{C2}), i.e., for the ages 11–15.

The survey in the Barents Sea has been conducted annually in February since 1981 and covers the demersal fishes in the Barents Sea, which includes the feeding grounds and mainly the immature part of the Northeast Arctic cod stock (Jakobsen et al. 1997). In the routine assessments, a swept area estimate is used (ICES 2004), i.e., the estimate is integrated to an absolute estimate for the surveyed area (Jakobsen et al. 1997). Because the surveyed area may vary from year to year, we chose to use the mean catch per trawled distance as an index of abundance instead of the reported swept area index (see Appendix A) to minimize the potential effect of year-to-year variations in surveyed area. We considered cod that were 3–9 years old.

Estimation of parameters

Catch-at-age data do not, in general, contain information about natural mortality, although it can estimate the total mortality ($Z_{a,y}$) under certain conditions (see, e.g., Quinn and Deriso 1999, p. 318). Indices of abundance can estimate the age-specific total mortality if the relative differences in age-specific catchability are known (substituting the mean value, eq. 6, into eq. 1 and solving with respect to $Z_{a,y}$ yields

$$\begin{aligned} Z_{a,y} &= \log(N_{a,y}/N_{a+1,y+1}) \\ &= \log([q_{a+1}E(I_{a,y})]/[q_a E(I_{a+1,y+1})]) \end{aligned}$$

If the relative differences in catchabilities are unknown, then the estimates of total mortality from surveys will be confounded with estimates of catchabilities. With a similar argument for the age-specific total mortalities, the age-specific catchabilities in the surveys could be estimated. If there is a lack of knowledge of either of the two, then the prospects of estimating age-specific natural mortality is poor without any additional information or, alternatively, using strong prior information on the parameters.

The set of parameters to be estimated is $(N_{a_{\min},1}, \ldots, N_{a_{\min},Y}, N_{a_{\min}+1,1}, \ldots, N_{A,1}, m_{a_{\min}}, \ldots, m_A, \sigma^2_M, \sigma^2_Z, f_{a_{\min}}, \ldots, f_{A^*}, \sigma^2_I, \sigma^2_I)$ for the population dynamical process, σ^2_{C1} and σ^2_{C2} for the observed catches, and $(q_{a^I_{\min}}, \ldots, q_{A^{(I)}}, \sigma^2_I)$ for the abundance indices from the survey. We chose vague prior distributions for most parameters, whereas the prior distributions for the parameters describing the natural mortality and the variance in the random walk (eq. 4) were chosen more informatively (see Appendix B for details).

To sample from the posteriors, we used the software package WinBUGS (Spiegelhalter et al. 1996) that uses MCMC methods to simulate the conditional distributions. Because this is an iterative simulation procedure that does not sample from the correct posteriors until convergence is reached, a certain number of iterates must be discarded before keeping the samples ("burn-in"). A variety of methods for assessing the convergence do exist (see, for example, Gelman et al. 2004), whereas we have examined this by starting three parallel Markov chains, with different (overdispersed) starting points, and applied the scale reduction factor (Gelman et al. 2004) together with visual inspections of the chains. The scale reduction factor measures the variability within and between the sequences and declines to 1 as the number of iterations approaches infinity. It is recommended that this factor should be smaller than 1.1 for all parameters for accepting convergence. Because samples from Markov chains per definition depend on earlier samples in the sequence, i.e., they are autocorrelated, it is necessary to run the Markov chains for a sufficiently long time to cover the entire posterior distribution. To reduce the number of iterations necessary to capture the entire distribution, it is therefore useful to thin the sequence by keeping only every kth iterate after convergence is obtained.

As expected, fitting the full model was not successful, which was evident from the lack of convergence of the Markov chains for the parameters, such that we were not able to estimate the posterior distributions of the parameters without using more informative priors on the age-specific

© 2007 NRC Canada

mortalities. With the modification that natural mortality for ages above the maximum age available in the surveys (9 years) was equal to that below this age, that is, $M_{a,y} = M_{9,y}$ for $a \geq 9$, the full model was fitted successfully regarding convergence of the parameters. However, the age-specific component of the natural mortality (m_a) and the age-specific selection in the fishing mortality (f_a) were highly correlated (close to 1), indicating that the model was overparameterised. Fitting the model to simulated data sets (see next section) confirmed the problems in separating those parameters. The correlations between f_a and the age-specific catchability in the surveys (q_a) and m_a and q_a were also rather high. To overcome problems with identifiability of the parameters, natural mortality was set as known and constant 0.2 for ages 10 years or older. For ages below 10 years old, the mean log natural mortality was set as constant, that is, $m_a = m$ for $3 \leq a \leq 9$.

For each Markov chain, we found that a burn-in of 20 000 iterates ensured convergence, and by keeping 334 samples of each parameter with a thinning of 60, yielding a total of 1002 samples per parameter, the posterior distributions was sufficiently well approximated, i.e., the percentiles in the distributions were stable. The total number of samples per parameter was then 120 120.

Simulation from the model

To evaluate the model fit and parameter estimates, we simulated both the population process and observations using the estimated parameters, analogous to ordinary parametric bootstrapping, forming replicates of both the population process and data. Each replicate of population and observations was obtained as follows. The population process was simulated with the same numbers of ages as the original data. It is necessary to generate the recruits each year. The model does not include a model for the recruits, which are considered as parameters. However, a stock–recruitment relationship is implicitly given as the spawning stock size can be estimated using the proportion mature at age. In this way we model the stock–recruitment relationship as the Ricker function, i.e.,

$$N_{a_{min},y} = \alpha S_{y-a_{min}} \exp(-\beta S_{y-a_{min}}) \exp(\varepsilon'_{a,y})$$

where α and β are positive parameters, S_y is the spawning stock size in year y, and the $\varepsilon'_{a,y}$s are iid normally distributed with zero mean and variance σ^2_1. The function was fitted to the posterior samples of the spawning stock size and the corresponding posterior samples of the recruits 3 years later using maximum likelihood. Each posterior sample of spawning stock size is estimated by the sum of the estimated number of mature individuals in each age group. We used the average of the annual estimates of proportion mature at age reported by ICES (2004) for the time period 1985–2003 to estimate the number of mature individuals, i.e., the proportions are constant for each age. The number of recruits was simulated by sampling at random from the residuals of the fit. To start the simulation process, it is necessary to initialise the process by setting the spawning stock size in some years in order to generate all initial values (see The population dynamical model), and we chose the initialisation period equal to the maximum age, i.e., the number of years needed for the process to generate the spawning stock itself. We used the average estimated spawning stock size as the initial value for the initialisation period.

The fishing and natural mortalities were simulated by their specified stochastic processes (eqs. 2–4) using the parameter estimates obtained by fitting the model to the data. We notice that even though we set the natural mortality as constant at 0.2 for the ages not included in the survey data when fitting the model to the observations, we simulate natural mortality according to its process (eq. 2) for all ages in the population. The random walk, modelling the log fishing effort (eq. 4), is not a stationary process and may thus produce unrealistically low or high fishing mortality rates over longer periods of time. To avoid this, we kept the fishing effort constant at 0 during the initialisation period and simulated as a random walk thereafter according to eq. 4.

Corresponding observations were then generated by simulating from the observation models (eqs. 5 and 6). The dimensions of simulated data sets are kept the same as those of the original data sets.

Results

The correlation among posterior samples of some of the parameters is shown (Fig. 1). The f_as are positively correlated with the q_as, decreasing with age and difference in age (the largest correlation is 0.88 for age 3), whereas both are negatively correlated with m, also with an absolute value that decreases with age (Fig. 1). Both m and σ^2_M and m and σ^2_1 are negatively correlated (–0.31 and –0.58, respectively). The correlation between the other parameters is weaker.

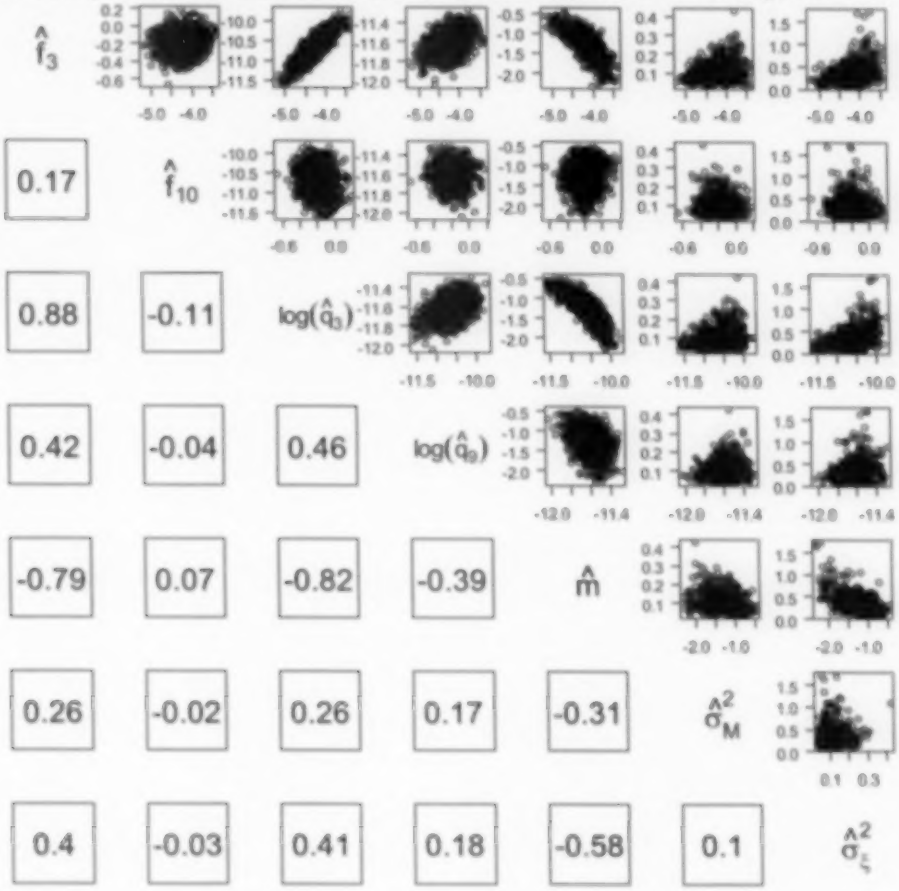
Estimating stock size and natural mortality

The estimates of the total population sizes for the time period 1985–2003 are shown (Fig. 2). The estimates showed rather large temporal fluctuations, with a mean ranging from more than five billion individuals in 1986 to below one billion in 1990. The estimated mean total population sizes are higher, on average, than those obtained by the Arctic Fisheries Working Group (AWFG; Fig. 2). Although the patterns in the temporal fluctuations were similar, the range of variation was smaller in the AFWG series (CV = 0.26) than in the corresponding mean for estimated population size (CV = 0.42). The annual variation in ECV in the estimates of the total population size ranged from 20% to 40% (i.e., the standard deviation of the posterior divided by the mean over the time series), but each posterior trajectory of total stock size was very similar (the average correlation among all pairs of samples of total stock size trajectories is 0.90).

The estimated mean annual natural mortality rate for the ages 3–9 years also showed large temporal fluctuations, from below 0.2 to above 0.7, and is on average 0.36, although the estimates were rather imprecise (Fig. 3*a*). The component accounting for the majority of the fluctuations in natural mortality was caused by year-to-year variability, giving a high correlation within years (between 0.5 and 0.9; Table 1) for ages 3–9 years.

© 2007 NRC Canada

Fig. 1. Pairwise plots of parameter estimates for log selectivity in fishing mortality for ages 3 and 10 ($\hat{f}_3$ and $\hat{f}_{10}$, respectively), log catchability in the survey for ages 3 and 9 ($\hat{q}_3$ and $\hat{q}_9$, respectively), log mean natural mortality (m), and the variance components in natural mortality ($\hat{\sigma}^2_M$ and $\hat{\sigma}^2_\xi$). The corresponding linear correlation coefficient is given in the lower diagonal.



Age-specific dynamics

The ages 5–10 years comprised the major proportion of the total annual landings of cod in the Barents Sea (ICES 2004). The mean annual fishing mortality for ages 5–10 years shows the same temporal dynamics as the one from the AFWG (ICES 2004) (Fig. 3*b*). The temporal variability in these estimates is slightly smaller than that from the estimates from the AFWG. The estimated variance component for the process for the fishing mortality σ^2_f (in eq. 3) is small (Table 1); thus, the deviation from the additive model for the log fishing mortality is rather small. Also the estimated effort employed by the fishing fleet changes relatively little from year to year, shown by a low estimate of the variance in the random walk for the log fishing effort (σ^2_f in eq. 4; Table 1).

The estimates of the variances for the input data correspond to an ECV of approximately 40% for the indices from the Barents Sea, 30% for the catch at age (ages below 11 years), and above 100% for catch at age for fish older than 11 years. Thus, the data for ages below 11 years seem relatively precise, whereas the catch data for older fish are very imprecise, i.e., contain little information about the fluctuations in abundance for these ages.

Stock size and recruitment

The relationship between spawning stock size and recruitment was fitted by the Ricker recruitment function. Large variation was found around the estimated stock–recruitment relationship (Fig. 4) because of the variability in the fitted values. The fit was used in the simulations to generate recruits given the spawning stock and did reproduce both the average abundance of the total stock and the estimated temporal fluctuations for the Northeast Arctic cod.

Simulations of fluctuations in stock size

By simulations it is evident that the fitted model estimates the total stock size very well (Fig. 5*a*), although the large stock sizes are slightly overestimated and small stock sizes are underestimated. By tracing the age-specific stock size, the model underestimated the actual number of recruits by an average of 15% (not shown). The magnitude of the bias decreased with age, resulting in overestimating the number

© 2007 NRC Canada

Fig. 2. The estimated total population size ($\times 10^{-6}$) of Northeast Arctic cod (*Gadus morhua*) age 3 years and older for the years 1985–2004 (solid line). The uncertainties in the estimates are shown by the 50% (dotted lines) and 95% (broken lines) credibility sets. The bold broken line represents the estimates from the Arctic Fisheries Working Group (ICES 2004).

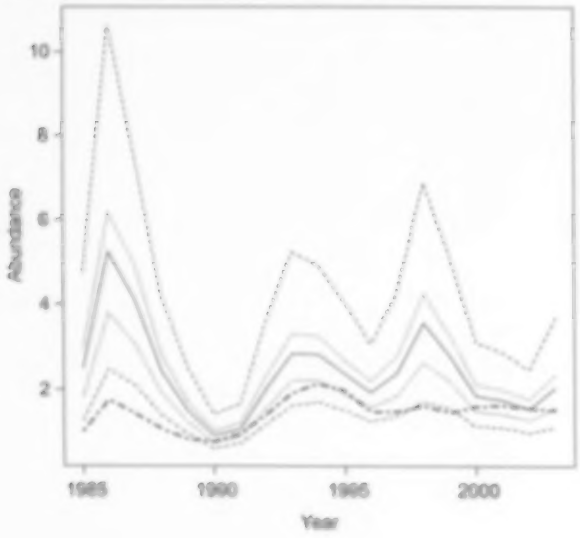
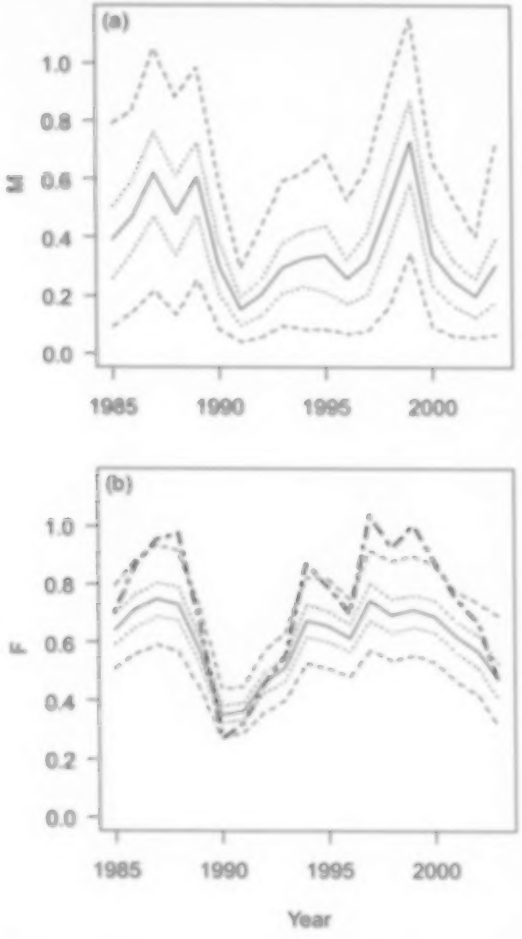


Fig. 3. The estimated annual mortality rates for the Northeast Arctic cod (*Gadus morhua*) stock for the years 1985–2003. (*a*) The estimated mean natural mortality rate for ages 3–9 years (solid line). The uncertainties in the estimates are shown by the 50% (dotted lines) and 95% (broken lines) credibility sets. (*b*) The estimated mean fishing mortality rate for ages 5–10 years (solid line). The uncertainties in the estimates are shown by the 50% (dotted lines) and 95% (dashed lines) credibility sets. The bold broken line represents the estimates given by the Arctic Fisheries Working Group (ICES 2004).



of fish in the oldest age class by about 15%. This is in agreement with Fig. 5*a* because a large total stock is dominated by young fish. Each replicate of process and corresponding data yields a correlation between the simulated time series of total population size and the corresponding estimated total population size. The mean correlation was 0.93 (95% credibility set is (0.79, 0.99)) (slightly higher than the correlation between the estimated and corresponding simulated total stock size independent of age as shown in Fig. 5*a*), showing that a high degree of the temporal fluctuations is captured by the model. The mean of the mean annual log natural mortality and the year-to-year variability (σ_ε^2) were underestimated (Table 1; Fig. 5*b*). Accordingly, the fluctuation in natural mortality was underestimated, resulting in overestimation of small values with a bias that decreases to underestimate large values of natural mortality (Fig. 5*b*). The variance component σ_M^2 in natural mortality was overestimated, and because the unconditional mean is given by $\exp(-m + 0.5(\sigma_M^2 + \sigma_\varepsilon^2))$, the mean of the natural mortality rate was underestimated by 20%. The level of fishing mortality independent of age was also very well estimated by the model with little bias (Fig. 5*c*). The age-specific fishing mortality was overestimated but the bias decreased with age (not shown). The relative bias for the mean fishing mortality for the ages 5–10 was about 10%.

The simulations showed that the 95% credibility sets for the estimated natural mortality and fishing mortality across ages and years both had an estimated coverage of 96.4%. Similarly, the estimated coverage for the estimated population sizes across ages and years was 95%.

We examined the ability of the model to estimate fluctuations in natural mortality outside the range of the estimated parameters for Northeast Arctic cod. In all cases we set the natural mortality within a year constant across ages, and all variability is due to temporal fluctuations. We considered five different scenarios: (*i*) constant natural mortality, but at different levels, (*ii*) linear increase in natural mortality over time, (*iii*) linear decrease in natural mortality, (*iv*) cyclic variations in natural mortality over time, and (*v*) stochastic variations in M independent between years according to a uniform distribution $U(a, b)$ for different a and b. All other population processes and corresponding observations were simulated as previously described.

In general, the results are in agreement with the previous simulations. We found that small values of natural mortality

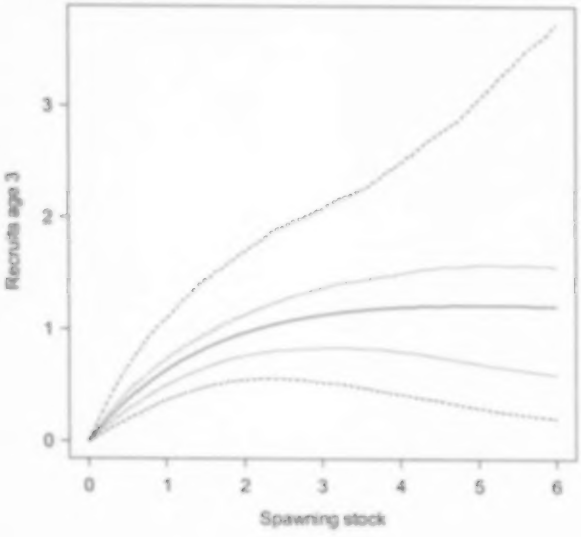
© 2007 NRC Canada

Table 1. Estimated population parameters for the model fitted to Northeast Arctic cod (see text): the mean, median, and 2.5% and 97.5% percentiles from the posterior distributions and the mean relative bias.

Parameter	Mean	Median	2.5% percentile	97.5% percentile	Mean relative bias (%)
m	−1.216	−1.171	−1.983	−0.689	−40.0
σ^2_M	0.097	0.086	0.045	0.212	48.7
σ^2_ε	0.311	0.266	0.103	0.825	−31.6
r_M	0.734	0.747	0.455	0.911	−23.3
σ^2_I	0.143	0.141	0.101	0.194	11.5
$\sigma^2_{C,1}$	0.073	0.072	0.050	0.106	33.8
$\sigma^2_{C,2}$	0.758	0.749	0.547	1.032	−6.5
σ^2_F	0.039	0.038	0.025	0.057	40.9
σ^2_δ	0.035	0.031	0.014	0.077	−6.4

Note: Mean log natural mortality, m; variance in log natural mortality, σ^2_M; covariance in log natural mortality, σ^2_ε; correlation within year in log natural mortality, r_M; variance in log index from the Barents Sea, σ^2_I; variance in log catch at age for ages 3–10 years, $\sigma^2_{C,1}$; variance in log catch at age for ages 11–15 years, $\sigma^2_{C,2}$; variance in log fishing mortality, σ^2_F; and variance in the random walk for the log effort in fishing mortality, σ^2_δ. The estimated relative bias (in percent) for the parameters is shown in the rightmost column and is based on refitting the models to 100 replicates of simulated data (see Fig. 5).

Fig. 4. The estimated mean recruitment ($\times 10^{-9}$) of 3-year-old Northeast Arctic cod (*Gadus morhua*) as a function of the estimated spawning stock three years earlier. The relationship is estimated by fitting the Ricker function to the posterior samples of the spawning stock and recruits. The solid line shows the resulting mean, and the dotted and broken lines are the corresponding 50% and 95% credibility sets for the mean recruitment, respectively.



were overestimated, whereas large Ms were underestimated, and vice versa. Thus, there is a consistent relationship between bias in M and bias in stock size. The relationship between bias in M and F is less consistent, although there seem to be a tendency for a negative relationship between the two, i.e., if M is underestimated, then F is overestimated. If considering the estimated temporal fluctuations compared with the simulated fluctuations rather than the absolute differences, we find that in all instances the estimated fluctuations in abundance are highly correlated with the simulated fluctuations (Table 2). The estimated expected correlation in the different scenarios ranges from 0.90 to 0.99, in which the correlation decreases with increasing variability in M. If the variability in temporal variations is sufficiently high, the estimated temporal variability is in close agreement with the simulated variability, although the variability is underestimated. The fluctuations in fishing mortality are in general better captured than the fluctuations in the natural mortality (higher correlations), and also here the temporal variability is underestimated, although to a lesser degree than that of the natural mortality.

Discussion

In this paper we have shown that it is possible to separate some elements of the stochastic natural and fisheries mortality by using catch-at-age data and survey indices. The major finding is that we can estimate the temporal dynamics in natural mortality, but that it is difficult to estimate the absolute value of this parameter. Also, fluctuations in stock size of Arctic cod in the Barents Sea were found to be closely related to temporal variation in natural mortality. Not surprisingly, the assessment of fish stock size would hence strongly benefit from more precise knowledge about the natural mortality.

It was not possible to estimate all parameters in the model without using any information in the prior distribution for some of the parameters (e.g., uniform prior distribution over a wide range). More specifically, the parameters most difficult to estimate were the variances in log natural mortality (σ^2_M) and log fishing mortality (σ^2_F). We therefore had to carefully select appropriate prior distributions restricting the values to a realistic range. This was also the case for the mean log natural mortality (m), although to a lesser extent than the former. We notice that the posterior distributions of the parameters do not reproduce the respective prior distribu-

© 2007 NRC Canada

Fig. 5. Comparisons of the simulated population process using the parameter estimates obtained by fitting the model to data for Northeast Arctic cod (*Gadus morhua*) with the estimates found by refitting the model to the corresponding simulated data for (*a*) annual total population size (N), (*b*) mean annual natural mortality ages 3–9 years (M), and (*c*) annual mean fishing mortality for ages 5–10 years (F). The solid lines are nonparametric fits of the data points to display the trend, the broken lines are the one-to-one mapping to illustrate deviation from the actual values, and the respective correlations (r) are shown in the figures. The results are based on refitting the model to 100 replicates of data sets. Each point represent an annual estimated mean versus the corresponding simulated mean, and each simulated data set is of length 19 years, thus each figure includes 1900 points.

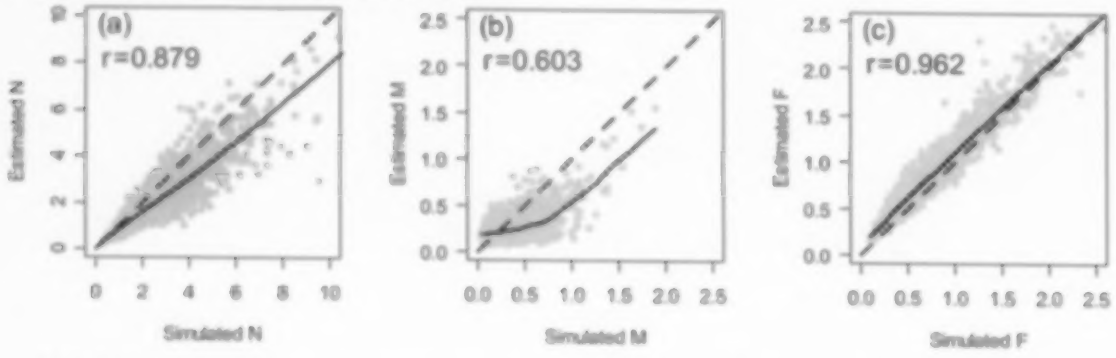


Table 2. Estimated mean correlation between temporal dynamics in estimated and simulated total stock size ($\mathrm{cor}(\hat{N},N)$), natural mortality ($\mathrm{cor}(\hat{M},M)$), and average annual fishing mortality for ages 5–10 years ($\mathrm{cor}(\hat{F},F)$) and estimated mean ratio of coefficient of variation (CV) of fluctuations in estimated to simulated CV of total stock size ($\mathrm{CV}(\hat{N})/\mathrm{CV}(N)$), natural mortality ($\mathrm{CV}(\hat{M})/\mathrm{CV}(M)$), and fishing mortality $\mathrm{CV}(\hat{F})/\mathrm{CV}(F)$ for different scenarios of simulated natural mortality.

Scenario for natural mortality	$\mathrm{cor}(\hat{N},N)$	$\mathrm{cor}(\hat{M},M)$	$\mathrm{cor}(\hat{F},F)$	$\frac{\mathrm{CV}(\hat{N})}{\mathrm{CV}(N)}$	$\frac{\mathrm{CV}(\hat{M})}{\mathrm{CV}(M)}$	$\frac{\mathrm{CV}(\hat{F})}{\mathrm{CV}(F)}$
$M_{a,y} = M = 0.01$	0.98	—	0.93	1.04	—	0.98
$M_{a,y} = M = 0.2$	0.98	—	0.94	1.01	—	0.89
$M_{a,y} = M = 0.5$	0.95	—	0.94	0.97	—	0.78
$M_{a,y} = M = 0.7$	0.98	—	0.92	1.02	—	0.68
$M_{a,y} = M = 1.0$	0.99	—	0.91	0.97	—	0.64
$M_{a,y} = M_y = (1/18)(y - 1)$	0.95	0.84	0.72	1.13	0.85	0.91
$M_{a,y} = M_y = (1/18)(19 - y)$	0.98	0.84	0.89	0.93	0.84	0.96
$M_{a,y} = M_y = 0.5(\sin(\pi y/4) + 1)$	0.90	0.91	0.81	1.16	0.69	0.90
$M_{a,y} = M_y = 0.25(\sin(\pi y/8) + 1)$	0.92	0.79	0.88	1.04	0.67	1.08
$M_{a,y} = M_y - U(0,1)$	0.93	0.72	0.90	0.96	0.55	0.82
$M_{a,y} = M_y - U(0,0.5)$	0.93	0.49	0.92	0.98	0.31	0.93
$M_{a,y} = M_y - U(0.1,0.3)$	0.97	0.17	0.94	1.03	0.38	0.91

Note: Scenarios considered: constant natural mortality at different levels (rows 1–5); systematic increase in natural mortality from 0 to 1 over 19 years (row 6); systematic decrease in natural mortality from 1 to 0 over 19 years (row 7); cyclic variations at different levels and with different periods (rows 8 and 9); stochastic variations in natural mortality independent between years according to a uniform distribution $U(a,b)$ for different a and b (rows 10–12). The other processes and data were simulated as for the Northeast Arctic cod case, and the same model was fitted to all simulated data sets (see text). The mean correlation and ratio CV was estimated as the average over 20 replicates of simulated process and data and the corresponding estimated process.

tions, and thus, there is some information in the data for those parameters, although the signal is weak (particularly for σ_M^2). This suggests that later studies should explore the strength of the signal for these parameters to derive appropriate priors beyond the method we have chosen to derive these priors. The other parameters in the model were not sensitive to the choice of prior distribution provided that the distribution was restricted to reasonable intervals. One of the strengths of the Bayesian approach is that it makes it possible to reflect our prior knowledge by a distribution, instead of setting unknown parameters in a model as constants. In situations in which some parameters in a model depend on the prior distribution, the sensitivity of the exact choice of prior distribution should be investigated (e.g., mean, variance, and type of distribution). Such detailed and systematic analysis is not done here and should be subject for further study. However, preliminary results suggest that the estimated temporal dynamics is not sensitive to other choices of reasonably chosen prior distributions but may cause changes in the estimated mean log natural mortality, which again affects the perceived absolute population size.

By using data from the Northeast Arctic cod stock, we have shown that this model type can estimate the fluctuations in population size, as well as separate the stochastic

© 2007 NRC Canada

components in mortality provided that we accept to use some prior information about some specific parameters. The precision in the estimates of total stock size (ECV is 20%–40%) appeared comparable to reported figures based on similar methods (Lewy and Nielsen 2003). Assuming that the natural mortality was a known constant or, analogously, choosing highly informative prior distributions for these parameters resulted in a dynamical pattern more closely related to the estimates provided by the AFWG. Obviously, the estimates will be less uncertain (ECV between 5% and 20%) if one assumes that the natural mortality is known. However, the samples from the posterior distribution of stock size – time series were highly correlated (the average correlation among all pairs of samples of stock size trajectories is 0.90), i.e., the dynamical pattern is nearly identical, but the absolute stock size differs. This indicates that the estimates are uncertain as a result of the same factor. Thus, the uncertainty in the estimated temporal fluctuations is lower than it appears in Fig. 2. Accordingly, there was a high correlation between the estimated abundance and the corresponding simulated data. Thus, gathering data for more precise estimates of natural mortality will greatly improve the precision in the estimates of the stock size.

The estimates of stock sizes and fishing mortalities are comparable with existing estimates provided by existing methods (ICES 2004), although this method shows different levels of variability. The estimates provided by the AFWG in 2004 for the population size and fishing mortality are in some years outside the estimated 95% credibility sets, but because these estimates do not provide any level of uncertainty, it is difficult to judge whether they are significantly different or not.

Our simulations revealed that our method gave poor estimates of the level of natural mortality that resulted in uncertain estimates of the total population size. This indicates that information about natural mortality obtained from other sources such as capture–recapture studies must be included in the model to obtain reliable estimates of total population size. Another possibility is to use a more strict Bayesian approach to put stronger constraints on the variation in the prior distributions of the natural mortality (m). Our method gives different estimates of total population size than by assuming a constant natural mortality of 0.2 with an additional mortality caused by cannibalism as employed by the AFWG (ICES 2004) for the cod. The resulting temporal dynamics in natural mortality is in poor agreement with the estimates found here, and we estimate higher natural mortality (mean of 0.36) with larger temporal fluctuations. This is the main reason for estimating the stock size to a higher level with larger temporal variability than the official estimates. Fortunately, our method seems to estimate annual fluctuations in the natural mortality quite well, enabling estimation of temporal fluctuations in stock size quite well. The high correlation between ages within year in natural mortality suggests that the major component causing fluctuations in this mortality is year-to-year variability. Empirical evidence from a variety of taxa suggests that temporal variation in life history traits with high sensitivities (Caswell 2001), such as adult survival rates in relatively long-lived species such as the Arctic cod, should be small (Pfister 1998; Sæther and Bakke 2000; Gaillard and Yoccoz 2003). Our findings of a large variability in natural mortality between years then may be primarily caused by discarding and black landings (ICES (2005), Myers et al. (1997), and Zeller and Pauly (2005) and references therein), or alternatively it may be caused by variability in rates of predation by sea mammals.

Another problem was to obtain reliable estimates of the variance in natural mortality (σ^2_M) because of a negative relationship ($r = -0.31$) between σ^2_M and m. This may be caused by variation in m that will induce fluctuations in age structure that will give autocorrelations in the process (Lande et al. 2006) and hence affect the estimates of σ^2_M.

In comparing the estimated precision levels in the catch data and the indices from the survey, it is apparent that, within this model, the catch data for the ages 3–10 years are only slightly more precise than those of the survey, whereas the catch data for the ages 11–15 years were very imprecise. This shows that the parameter estimates rely mainly on the data for ages up to 10 years and little on the ages above. Hence, leaving out the data for the oldest fish would probably not result in significant changes for the estimates. The estimated precision levels of the input data contain both sampling error and process error and are confounded within this model. If the sampling error is known beforehand, then subject to some assumptions, the total error is the sum of sampling and process error (Lewy and Nielsen 2003). In the present paper, the sampling error (not shown) for the Norwegian catches were, on average, larger than the process error for all catches (estimated by modelling the sampling schemes; see Aanes and Pennington (2003) and Hirst et al. (2004, 2005)). However, because of the lack of data from other than Norwegian catches, we can not exclude the possibility that process error may be of more significant importance than the error related to sampling (see Lewy and Nielsen (2003) for an example on North Sea plaice).

We assumed that the errors in the input data were independent and identically distributed across ages and years. The effect of violating this assumption is not examined, but in general, this could cause biases in the estimates, with a magnitude depending on the magnitude of correlation at age within year and the relative differences in precision between ages. The magnitude of this potential problem could be explored as in Myers and Cadigan (1995*a*; 1995*b*), who allowed the survey indices to be correlated within years and quantified by simulations similar to what is done in this paper, or alternatively one could try to estimate the entire covariance matrix. Estimating the entire covariance matrix would probably lead to overparameterisation of the model, and it would therefore be advantageous to parameterise the covariance matrices with as few parameters as possible to avoid this problem. Lillegård et al. (2005) examined the process noise for the spawning stock of the Norwegian spring spawning herring and found only a very small temporal correlation, which was mainly generated by a small temporal correlation in temperature.

The model was able to reproduce simulated data. Although some parameter estimates were biased, the estimates of the most important figures — the total stock size, natural mortality, and fishing mortality — were very accurate with relatively small biases. The reported biases for stock size and fishing mortality agree with the estimated biases found by Lewy and Nielsen (2003), and the estimated coverage of

© 2007 NRC Canada

the credibility sets was close to the true probability. For simulation purposes, we fitted the stock–recruitment function to the posterior estimates of spawning stock size and recruits. The fit showed large variability around the mean, which could have been caused by factors such as process error and (or) environmental noise (Mann and Drinkwater (1994), Ottersen et al. (2001), and Myers (2002) and references therein). The fitted function is, however, in agreement with other fits (e.g., Myers et al. 1996, 1997).

Acknowledgements

The authors thank M. Lillegård, D. Skagen, A. Salthaug, and A. Aglen for valuable discussions. This paper is funded by Institute of Marine Research (Bergen, Norway) (IMR), the Norwegian University of Science and Technology (NTNU), the Norwegian Polar Institute (NPI), and the Norwegian Research Council (NFR) through the NORKLIMA programme. The authors also thank two anonymous referees and the Editors for valuable comments.

References

Aanes, S., and Pennington, M. 2003. On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. ICES J. Mar. Sci. **60**: 297–303.

Caswell, H. 2001. Matrix population models. 2nd ed. Sinauer Associates, Inc., Sunderland, Massachusetts.

Clark, J.S. 2005. Why environmental scientists are becoming Bayesians. Ecol. Lett. **8**: 2–14.

Clark, J.S., and Bjørnstad, O.N. 2004. Population time series: process variability, observation errors, missing values, lags, and hidden states. Ecology, **85**: 3140–3150.

Dennis, B., Ponciano, J.M., Lele, S.R., Taper, M.L., and Staples, D.F. 2006. Estimating density dependence, process noise, and observation error. Ecol. Monogr. **76**: 323–341.

de Valpine, P., and Hastings, A. 2002. Fitting population models incorporating process noise and observation error. Ecol. Monogr. **72**: 57–76.

de Valpine, P., and Hilborn, R. 2005. State–space likelihoods for nonlinear fisheries time-series. Can. J. Fish. Aquat. Sci. **62**: 1937–1952.

Freckleton, R.P., Watkinson, A.R., Green, R.E., and Sutherland, W.J. 2006. Census error and the detection of density-dependence. J. Anim. Ecol. **75**: 837–851.

Gaillard, J.M., and Yoccoz, N.G. 2003. Temporal variation in survival of mammals: a case of environmental canalization? Ecology, **84**: 3294–3306.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 2004. Bayesian data analysis. Chapman and Hall, London, UK.

Gilks, W.R., and Wild, P. 1992. Adaptive rejection sampling for Gibbs sampling. Appl. Stat. **41**: 337–348.

Gudmundsson, G. 1994. Time series analysis of catch at age observations. J. R. Stat. Soc. Ser. C, Appl. Stat. **43**: 117–126.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York.

Hirst, D., Aanes, S., Storvik, G., and Tvete, I.F. 2004. Estimating catch at age from market sampling data using a Bayesian hierarchical model. J. R. Stat. Soc. Ser. C, Appl. Stat. **53**: 1–14.

Hirst, D., Storvik, G., Aldrin, M., Aanes, S., and Huseby, R.B. 2005. Estimating catch-at-age by combining data from different sources. Can. J. Fish. Aquat. Sci. **62**: 1377–1385.

International Council for the Exploration of the Sea. 2004. Report of the Arctic Fisheries Working Group, ICES Advisory Committee on Fishery Management. ICES CM 2004/ACFM:28.

International Council for the Exploration of the Sea. 2005. Report on the assessment of demersal stocks in the North Sea and Skagerrak, ICES Advisory Committee on Fishery Management. ICES CM 2005/ACFM:07.

Jakobsen, T., Korsbrekke, K., Mehl, S., and Nakken, O. 1997. Norwegian combined acoustic and bottom trawl surveys for demersal fish in the Barents Sea during winter. ICES CM 1997/Y:17.

Lande, R., Engen, S., and Sæther, B.-E. 2003. Stochastic population dynamics in ecology and conservation. Oxford University Press, New York.

Lande, R., Engen, S., Sæther, B.-E., and Coulson, T. 2006. Estimating density dependence from time series of population age structure. Am. Nat. **168**: 76–87.

Lewy, P., and Nielsen, A. 2003. Modelling stochastic fish stock dynamics using Markov chain Monte Carlo. ICES J. Mar. Sci. **60**: 743–752.

Lillegård, M., Engen, S., Sæther, B.-E., and Toresen, R. 2005. Harvesting strategies for Norwegian spring-spawning herring. Oikos, **110**: 567–577.

Ludwig, J., and Walters, C.J. 1989. A robust method for parameter estimation from catch and effort data. Can. J. fish. Aquat. Sci. **46**: 137–144.

Mann, K.H., and Drinkwater, K.F. 1994. Environmental influences of fish and shellfish productions in the Northwest Atlantic. Environ. Rev. **2**: 16–32.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Am. Fish. Soc. Symp. **6**: 8–48.

Millar, R.B., and Meyer, R. 2000. Bayesian state–space modeling of age-structured data: fitting a model is just the beginning. Can. J. Fish. Aquat. Sci. **57**: 43–50.

Myers, R.A. 2002. Recruitment: understanding density dependence in fish populations. *In* Handbook of fish biology and fisheries. Vol 1. Fish biology. *Edited by* P.J. Hart and J. Reynolds. Blackwell Publishing Ltd., Oxford, UK. pp. 123–148.

Myers, R.A., and Cadigan, N.G. 1995*a*. Statistical analysis of catch-at-age data with correlated errors. Can. J. Fish. Aquat. Sci. **52**: 1265–1273.

Myers, R.A., and Cadigan, N.G. 1995*b*. Was an increase in natural mortality responsible for the collapse of northern cod? Can. J. Fish. Aquat. Sci. **52**: 1274–1285.

Myers, R.A., Hutchings, J.A., and Barrowman, N.J. 1996. Hypotheses for the decline of cod in the North Atlantic. Mar. Ecol. Prog. Ser. **138**: 293–308.

Myers, R.A., Hutchings, J.A., and Barrowman, N.J. 1997. Why do fish stocks collapse? The example of cod in Atlantic Canada. Ecol. Appl. **7**: 91–106.

Ottersen, G., Planque, B., Belgrano, A., Post, E., Reid, P.C., and Stenseth, N.C. 2001. Ecological effects of the North Atlantic Oscillation. Oecologia, **128**: 1–14.

Patterson, K.R. 1999. Evaluating uncertainty in harvest control law catches using Bayesian Markov chain Monte Carlo virtual population analysis with adaptive rejection sampling and including structural uncertainty. Can. J. Fish. Aquat. Sci. **56**: 208–221.

Pfister, C.M. 1998. Patterns of variance in stage-structured populations: evolutionary predictions and ecological implications. Proc. Natl. Acad. Sci. U.S.A. **95**: 213–218.

Punt, A.E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish Biol. Fish. **7**: 35–63.

© 2007 NRC Canada

Quinn, T.J., and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York.
Richards, L.J., and Schnute, J.T. 1998. Modal complexity and catch-age analysis. Can. J. Fish. Aquat. Sci. **55**: 949–957.
Schnute, J.T. 1994. A general framework for developing sequential fisheries models. Can. J. Fish. Aquat. Sci. **51**: 1676–1688.
Schnute, J.T., and Richards, L.J. 1995. The influence of error on population estimates from catch-age models. Can. J. Fish. Aquat. Sci. **52**: 2063–2077.
Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. 2nd ed. MacMillian, New York.
Spiegelhalter, D.J., Thomas, A., Best, N.G., and Gilks, W.R. 1996. BUGS 0.5: Bayesian inference using Gibbs sampling, manual (version *ii*). MRC Biostatistics Unit, Cambridge, UK.
Sæther, B.-E., and Bakke, Ø. 2000. Avian life history variation and contribution of demographic traits to the population growth rate. Ecology, **81**: 642–653.
Sæther, B.-E., Lillegård, M., Grøtan, V., Filli, F., and Engen, S. 2007. Predicting fluctuations of re-introduced ibex populations: the importance of density-dependence, environmental stochasticity and uncertain population estimates. J. Anim. Ecol. **76**(2): 326–336.
Tierney, L. 1998. A note on the Metropolis Hastings algorithm for general state spaces. Ann. Appl. Prob. **8**: 1–9.
Walters, C.J., and Martell, S.J.D. 2004. Fisheries ecology and management. Princeton University Press, Princeton, New Jersey.
Zeller, D., and Pauly, D. 2005. Good news, bad news: global fisheries discards are declining, but so are total catches. Fish Fish. (Oxford), **6**: 156–159.

Appendix A. Survey indices.

Table A1. Stratified average catch in numbers per towed distance for ages 3–9 for the annual survey in February for demersal fishes in the Barents Sea (Jakobsen et al. 1997) for the years 1985–2003.

	Age						
Year	3	4	5	6	7	8	9
1985	27.04	19.74	3.33	1.46	0.54	0.10	0.01
1986	130.37	29.68	11.83	1.72	0.23	0.02	0.01
1987	43.91	68.32	9.56	3.13	0.31	0.11	0.00
1988	19.79	10.37	24.36	3.08	0.60	0.07	0.01
1989	4.90	9.51	5.48	13.74	1.42	0.17	0.02
1990	2.95	4.25	4.96	3.15	4.09	0.29	0.04
1991	3.60	2.65	2.35	2.20	1.23	1.27	0.06
1992	25.84	10.19	3.44	2.34	1.40	0.73	0.52
1992	29.12	18.56	8.48	1.83	0.91	0.49	0.26
1994	23.81	30.99	14.30	4.93	0.94	0.22	0.12
1995	18.91	17.00	20.14	6.20	1.47	0.22	0.05
1996	10.06	10.11	11.95	11.09	2.63	0.31	0.04
1997	18.10	6.83	8.07	7.19	3.89	0.71	0.13
1998	43.83	28.42	6.32	4.37	2.84	1.33	0.19
1999	21.98	23.16	10.34	2.33	1.13	0.49	0.11
2000	17.03	9.25	8.54	2.51	0.41	0.18	0.09
2001	14.01	14.66	7.57	3.76	0.85	0.10	0.05
2002	7.28	10.80	9.91	4.49	1.50	0.23	0.03
2003	23.35	10.57	10.86	8.65	2.40	0.59	0.12

Note: The data are stratified according to the actual spatial survey design the given year. The catches before 1989 are adjusted according to the change from Bobbins to Rockhopper gear using established correction factors (Godø and Sunnanå 1992).

Appendix B. Prior distributions.

The prior distributions for the parameters in the model described in Model are chosen as follows.

For the log recruits ($\log(N_{a_{min},y})$, $y = 1,...,Y$), the log initial values ($\log(N_{a,1})$, $a = a_{min}, ..., A$), the log catchabilities in the surveys ($\log(q_a)$, $a = a_{min}^{(I)}, ..., A^{(I)}$), and the log fishing selectivity ($\log(f_a)$, $a = a_{min}, ..., A^*$), we chose normally distributed priors with variance 1000, and means 10, 10, −5, and −1.5, respectively. The mean in the distributions are arbitrarily chosen, and the posterior distribution does not depend on these choices.

The estimates of fishing mortality given in ICES (2005) shows age-specific temporal variation in fishing mortality corresponding to a coefficient of variation of slightly above 0.2 for the time period and age groups considered in this paper. This corresponds to the variance in the lognormal distribution being ≈0.04. Therefore, we parameterise the inverse gamma for the variance in the random walk modelling the fishing effort σ_f^2 such that the mean is 0.04 and variance is 0.01, and consequently 95% of the probability mass of the prior distribution is between ≈0 and 0.17 (see Figs. B1*a* and B1*b*).

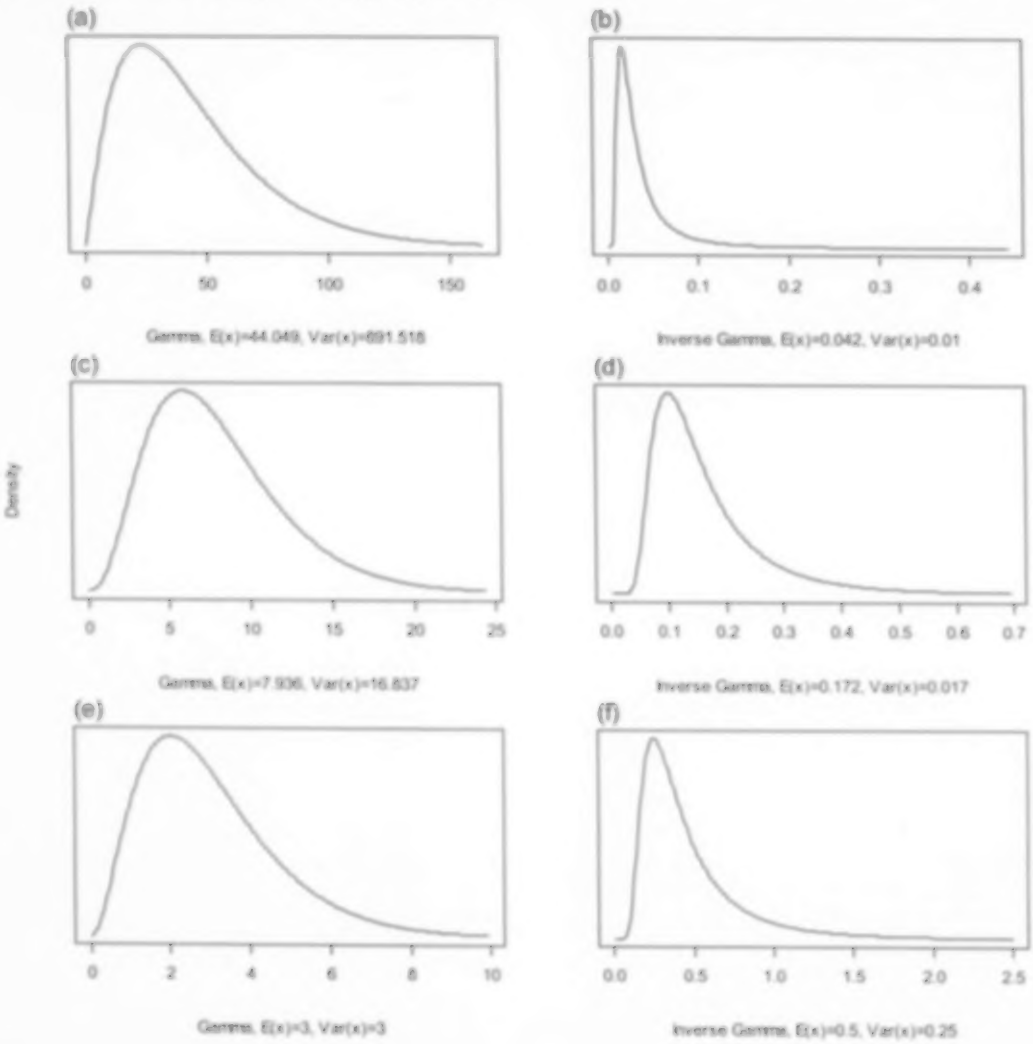
© 2007 NRC Canada

The prior for the random variability in fishing mortality (σ_f^2) is chosen the same as for the variance components in natural mortality (see rationale below).

The variance in natural mortality rarely exceeds 0.04 (Sæther et al., unpublished). Because the natural mortality is modelled as a lognormally distributed process, i.e., $M_{a,y} \sim \text{LN}(\mu, \sigma^2)$, the variance of M is given by $(E(M))^2 (e^{\sigma^2} - 1)$. Then, given that $E(M) = 0.2$, the variance of the mortality rate is constrained to $0.04 > 0.2^2 (e^{\sigma^2} - 1) \sim \sigma^2 < \log(2)$. To achieve this, we chose to parameterise the inverse gamma for both components in natural mortality, σ_ξ^2 and σ_M^2, such that the mean is 0.172 with variance 0.017. Then for each component, 95% of the probability mass of the prior distribution is in the interval (0.05, 0.5) (see Figs. B1*c* and B1*d*). This choice results in prior distribution of median M with 95% probability mass in the interval (0.03, 1.42).

The priors for the variance for the error distribution for the input data, $\sigma_{\varepsilon,1}^2$, $\sigma_{\varepsilon,2}^2$, and σ_I^2, were parameterised such that both the mean and standard deviation are 0.5, and consequently 95% of the probability mass of the error distribution is in the interval (0.02, 2) (see Figs. B1*e* and B1*f*).

Fig. B1. Prior distribution is gamma for the inverse variance or equivalently inverse gamma for the variance. (*a* and *b*) The prior distribution of the variance in the random walk modelling the year to year changes of log fishing mortality (σ_f^2); (*c* and *d*) the random variation in fishing mortality (σ_f^2) and the variance components for the natural mortality (σ_ξ^2 and σ_M^2); (*e* and *f*) the prior for the variance in the error distributions for the input data ($\sigma_{\varepsilon,1}^2$, $\sigma_{\varepsilon,2}^2$, and σ_I^2).



© 2007 NRC Canada

Movements of adult coho salmon (*Oncorhynchus kisutch*) during colonization of newly accessible habitat

Joseph H. Anderson and Thomas P. Quinn

Abstract: Pacific salmon (*Oncorhynchus* spp.) have repeatedly exploited new habitat following glacial recession and some artificial introductions, yet the initial process of colonization is poorly understood. Landsburg Diversion Dam on the Cedar River, Washington, excluded salmon from 33 km of habitat for over a century until it was modified to allow passage in 2003. Adult coho salmon (*Oncorhynchus kisutch*) were sampled as they entered the newly accessible habitat in the first 3 years and a subset received radio transmitters to assess spawning site selection and movement. Annual counts of coho colonists increased over time, and in 2 of 3 years, daily dam passage was positively correlated with river discharge. Contrary to our prediction that coho would spawn in tributaries, all identified spawning sites were in the mainstem Cedar River, though 38% of radio-tagged salmon entered a tributary at least temporarily. Females moved little within the new habitat (average = 5.8 km), whereas males moved extensively (average = 34.8 km), especially when females were scarce. The immediate use of the new habitat by colonists and their widespread movements suggest that exploration is an innate component of salmon breeding behavior, and restoring access to lost habitat merits prioritization as a conservation strategy.

Résumé : Après le retrait des glaciations et à la suite d'introductions artificielles, les saumons du Pacifique (*Oncorhynchus* spp.) ont exploité de façon répétée de nouveaux habitats; néanmoins, le processus initial de colonisation reste peu compris. Le barrage de dérivation Landburg sur la rivière Cedar, Washington, a exclu les saumons d'un habitat de 33 km pendant plus d'un siècle jusqu'à ce qu'il soit modifié en 2003 pour permettre leur passage. Les saumons coho (*Oncorhynchus kisutch*) ont été échantillonnés durant les 3 premières années après leur pénétration dans l'habitat nouvellement disponible; un sous-ensemble de ces poissons ont été munis de transmetteurs radio afin de déterminer leur choix de site de fraye et leurs déplacements. Les dénombrements annuels de saumons coho colonisateurs se sont accrus avec le temps et dans 2 des 3 années leur passage journalier au barrage était en corrélation positive avec le débit de la rivière. Contrairement à notre prédiction que les saumons coho frayeraient dans les tributaires, tous les sites de fraye identifiés se situent dans le cours principal de la rivière Cedar, bien que 38 % des saumons munis d'une étiquette radio ait pénétré au moins temporairement dans un tributaire. Les femelles se déplacement peu dans le nouvel habitat (moyenne = 5,8 km), alors que les mâles se déplacent beaucoup (moyenne = 34,8 km), particulièrement lorsque les femelles sont rares. L'utilisation immédiate du nouvel habitat par les colonisateurs et leurs déplacements étendus indiquent que l'exploration est une composante innée du comportement de fraye des saumons; il y a donc intérêt à restaurer l'accès aux habitats perdus comme stratégie prioritaire de conservation.

[Traduit par la Rédaction]

Introduction

Habitat colonization is critical to population biology, as a species' range reflects the countervailing forces of colonization and local extinction. Strict philopatry (i.e., homing to the natal site for reproduction) precludes exploration and possible colonization of new breeding areas. It is unclear whether colonization results from orientation failure or active exploration of new areas despite properly functioning homing mechanisms. Orientation failure may place colonists into new habitats over a range of distances from the source population. Alternatively, non-natal spawning sites with more favorable environmental conditions or reduced competition may attract breeders to explore and colonize new areas, though this process may operate over shorter distances than orientation failure. Indeed, a population must contain at least a small proportion of exploratory rather than philopatric spawners to ensure long-term population viability (LePage and Cury 1997). Exploratory behavior is of immediate significance to a population because homing by the offspring of colonists permits new breeding areas to become established in a single generation (Cury 1994). By describ-

Received 25 August 2006. Accepted 18 April 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 10 August 2007. J19503

J.H. Anderson[1] and T.P. Quinn. School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195, USA.
[1]Corresponding author (e-mail: joeander@u.washington.edu).

 doi:10.1139/F07-087 © 2007 NRC Canada

ing the behavior of colonizing individuals, we can understand mechanisms of population dispersal and its significance to population structure (Rieman and Dunham 2000).

Homing to the natal spawn site is a hallmark of salmonid fishes such as Pacific salmon (*Oncorhynchus* spp.), yet some "strays" survive to maturity and spawn away from their natal site (Quinn 1993). Homing, combined with generations of natural selection, leads to the evolution of traits advantageous to local conditions (Ricker 1972; Taylor 1991; Quinn 2005). However, strays are important to the evolutionary plasticity and geographic distribution of the species, as they allow for colonization of new habitat. Both postglacial history and experience with transplanted populations illustrate that salmon can exploit open habitat. Salmon expanded into areas made available by receding glaciers from northern Washington to southern Alaska and proliferated in a myriad of geologically young streams (McPhail and Lindsey 1986). This colonization process continues in regions of glacial recession such as Glacier Bay, Alaska, where salmon colonization follows development of complex stream habitat (Milner and Bailey 1989; Milner et al. 2000). Further evidence for the colonization potential of Pacific salmon comes from artificial introductions outside their native ranges, including pink salmon, *O. gorbuscha*, in the Great Lakes (e.g., Kwain 1987) and Chinook salmon, *O. tshawytscha*, in New Zealand (e.g., Quinn et al. 2001) and Argentina (Ciancio et al. 2005). In each case, salmon rapidly dispersed into nearby unoccupied streams and established self-sustaining populations. Unfortunately, the initial stages of these colonization events are shrouded in the past, and processes occurring at the onset of colonization are not well understood.

Furthermore, we know little about the factors promoting successful colonization. We postulate that success might depend on three behavioral processes: (*i*) entry into the new habitat, (*ii*) exploratory movements before breeding, and (*iii*) selection of spawning sites. Research on entry into new habitat should document the basic attributes of salmon colonists (e.g., abundance, seasonal timing, sex, body size, and origin, including whether the fish were produced naturally or in a hatchery) as population expansion proceeds through time and the environmental conditions affecting upriver movement into new areas. Hydrologic conditions might be particularly important, as upriver movement of salmonids is often triggered by increased discharge in small streams (Banks 1969). Although the effects of water temperature on upriver movement are variable (Banks 1969), salmonids avoid extremely warm areas during spawning migrations (Berman and Quinn 1991; Newell and Quinn 2005).

Colonization may result from a breakdown of olfaction and homing to the natal site (Dittman and Quinn 1996), from active choice of a non-natal breeding area, or from some combination of the two. Thus exploration and spawning site selection are important components of colonization behavior. Females need to locate suitable breeding sites because they alone prepare nests, so movement and spawning site selection are intimately linked. Coho salmon, *Oncorhynchus kisutch*, for example, tend to use small tributaries for breeding (Sandercock 1991; Montgomery et al. 1999; Quinn 2005), and densities of juvenile coho salmon are also typically highest in tributaries of larger rivers (Rosenfeld et al. 2000; Scarnecchia and Roper 2000), although some rear in mainstem sections of rivers (Peters 1996). During the early stages of colonization, densities will presumably be low relative to established populations, and female salmon may not be restricted by competition for spawning sites (e.g., van den Berghe and Gross 1989), or by homing to the natal site, and thus female movements may directly reflect choice of breeding habitat.

Movements of adult salmon vary according to stage of sexual maturation, as migrating salmon behave differently than spawning salmon. Migrating salmon can move rapidly (Keefer et al. 2004) and may show a "search phase" prior to spawning characterized by erratic upstream and downstream movements, presumably as they search for natal odors or select breeding habitat (Økland et al. 2001; Connor and Garcia 2006). In large rivers, these search movements cover long distances (Keefer et al. 2006). However, there have been fewer studies of adult salmon movements on the spawning grounds, and the results have been mixed. Movement by males may be determined by the distribution of ripe females in both space and time. Radio tracking revealed considerable movement by male coho salmon, likely in search of mates (Healey and Prince 1998), but spawning ground movements of male sockeye salmon (*Oncorhynchus nerka*) experiencing much higher intraspecific breeding densities were very limited (Hendry et al. 1995; Stewart et al. 2004; Rich et al. 2006).

The construction of a fish ladder by Seattle Public Utilities at Landsburg Diversion Dam on the Cedar River, Washington (Fig. 1), presented a rare opportunity to investigate the process of salmon colonization. The ladder, completed in fall 2003, made 33 km of suitable spawning habitat accessible to Chinook salmon, coho salmon, and steelhead trout (*Oncorhynchus mykiss*) that had been excluded since the dam was constructed in 1900. The newly available habitat includes the mainstem Cedar River and four tributaries, of which Rock Creek offers the majority of potential salmon habitat (accessible area with moderate gradient, woody debris, and a mixture of pools and riffles).

To investigate the process of colonization, adult coho salmon were sampled as they passed Landsburg Diversion Dam, and some were radio-tracked to characterize their movements. The study had three principal objectives. First, we recorded basic demographics of colonists entering the new habitat: abundance, timing, body size, and origin (hatchery or naturally spawned). Dam passage data were secondarily used to evaluate the proximate environmental stimuli for upriver movement into the newly accessible area. Second, we assessed reach-scale patterns in spawning site selection, hypothesizing that coho salmon would spawn primarily in tributaries, especially Rock Creek, rather than the mainstem Cedar River. Third, we determined the effects of river discharge, date, body size, and access to breeding opportunities on the movements of individual salmon in the newly accessible habitat. We compared the movements of females having access to a large number of potential breeding sites with those of males having access to a small number of possible mates. Our analysis tested the hypothesis that salmon movements would be inversely related to breeding opportu-

© 2007 NRC Canada

nity: males would be more mobile than females, and males would be more mobile when reproductively active females were scarce.

Materials and methods

Study site description

The Cedar River drains a 487 km^2 watershed in Washington State, flowing westward from the crest of the Cascade Mountains (Fig. 1) into the southern end of Lake Washington, and then to marine waters via a shipping canal. In 1900, the City of Seattle constructed Landsburg Diversion Dam, a low head, run-of-the-river facility, at river kilometre 35.1 to supply Seattle residents with drinking water. This dam blocked salmonid migration until passage facilities were completed in fall 2003, making available over 33 km of habitat in the mainstem Cedar River and tributaries. Much of the landscape is dominated by second-growth coniferous forest aged at least 60 years (Seattle Public Utilities 2000), and the City of Seattle now manages the area above Landsburg Diversion Dam as a de facto reserve, without development, recreation, or commercial logging. Coho salmon spawn naturally in the Cedar River and tributaries below Lansdsburg Diversion Dam, although no comprehensive population estimates were available. In contrast to the upper watershed, land use of the Cedar River watershed below the dam is dominated by suburban development.

Of the 33 km of new habitat now accessible to salmon, 20 km are located in the Cedar River between Landsburg Diversion Dam and lower Cedar Falls, a natural barrier to migration. The Cedar River has a relatively confined channel, and densities of instream large woody debris are lower than in similar streams in western Washington (Kiffney et al. 2002). The remaining 13 km of newly accessible habitat are found in the tributaries (Rock, Williams, Steele, and Taylor creeks). Only very short sections (<0.5 km) of Williams, Steele, and Taylor creeks are accessible to salmon before they are blocked by natural barriers. The majority of tributary habitat available to salmon is in Rock Creek, which is also the first tributary that a salmon would encounter swimming upriver from the Landsburg fish passage facility. For these reasons, much of the radio-tracking effort was directed to Rock Creek.

In addition to salmon, the Cedar River watershed below Cedar Falls contains resident fish typical of many streams in western Washington. Salmonid species include rainbow trout (*Oncorhynchus mykiss*), cutthroat trout (*Oncorhynchus clarkii*), and mountain whitefish (*Prosopium williamsoni*). In general, trout densities are much lower in the Cedar River than in rivers of comparable size in western Washington (Kiffney et al. 2002). Nonsalmonid fishes in Cedar River and tributaries include speckled dace (*Rhinichthys osculus*), brook lamprey (*Lampetra richardsoni*), and several sculpin species (*Cottus* spp.) that are numerically dominant.

Movement through the fish passage facility

Adult coho salmon were sampled as they moved upriver through the fish passage facility at Landsburg Diversion Dam. For the great majority of the run, a gate prevented salmon from accessing the new habitat unless they were handled. The facility was operated in this mode from before the run began until the run was deemed over (0–1 fish migrating upriver in 2 weeks). The date of passage, species, sex, and length were recorded for each fish. Fork length was measured in 2004 and 2005; total lengths measured in 2003 were converted to fork lengths with data from the nearby University of Washington hatchery using sex-specific linear regressions (males, $n = 28$, $p < 0.0001$, $r^2 = 0.997$; females, $n = 21$, $p < 0.0001$, $r^2 = 0.987$). The fish were also examined for the presence or absence of an adipose fin as an indication of natural spawning or hatchery origin, respectively. In the brood years that returned during this study, virtually all hatchery-produced coho salmon were adipose-clipped at the two hatcheries in the basin producing this species (Issaquah Creek and University of Washington). After the fish passage gate was opened late in each season, a few salmon might have passed upriver unsampled but an automatic camera indicated that unsampled fish were <5% of the total run in any year.

Fig. 1. Map of Cedar River watershed above Landsburg Diversion Dam (Washington, USA). Broken lines represent habitat that is inaccessible to salmon, as indicated on maps provided by Seattle Public Utilities. Shaded circles and two letter codes refer to the listening stations, where the size of the circle is not proportional to the range of the antennas.

Generalized least squares (GLS) regression models were used to determine the relationship between upriver movement past Landsburg Diversion Dam and both temperature and discharge, recorded at US Geological Survey gauges located within 3 km of the dam. The data used were from 1 week before the first upriver movement through the ladder to 1 week after the last upriver movement in each season ($n = 90$ days in 2003, 130 days in 2004, and 121 days in 2005). With adult counts as the response variable, regression explored four predictor variables: mean daily discharge (Q), change in discharge from the previous day (ΔQ_{t-1}), mean daily temperature (T), and change in temperature from the previous day (ΔT_{t-1}). Environmental and count data were collected on successive days, and a GLS approach considers such data as a times series (Trépanier et al. 1996). The model error structure was specified as a first-order autoregressive process and parameters were estimated via maxi-

© 2007 NRC Canada

mum likelihood (Ostrom 1990). For the GLS models, listed p values are from t tests for the significance of each predictor, and we also report Nagelkerke's (1991) adjusted r^2, which compares the log likelihood of the fitted model with the log likelihood of the null (intercept only) model.

Telemetry

Some of the coho salmon bypassing Landsburg Diversion Dam in 2003 and 2004 received radio transmitters via esophageal implant. The transmitters (model F1845, Advanced Telemetry Systems, Inc., Isanti, Minnesota) weighed 24 g in air, with a maximum battery life of 360 days. Movements of radio-tagged fish were determined by listening stations and mobile tracking. Listening stations at strategic locations along the mainstem Cedar River and tributaries continuously scanned for transmitters (Fig. 1) with a receiver–datalogger unit and three antennas positioned to indicate the direction of movement. Four such listening stations were located in the newly accessible habitat in fall 2003: the lower Rock Creek (LR) station was at the confluence with the mainstem; the middle Rock Creek (MR) station was 2.5 km upstream from the Cedar River; and the upper Rock Creek (UR) station was 4.6 km up Rock Creek from the Cedar River. The final station was at the confluence of the Cedar River and Taylor Creek (TA). In fall 2004, two new listening stations were added: one at Landsburg Dam (LD) and another on the middle of the mainstem Cedar River (MM), halfway between Rock and Taylor creeks. The initial distribution of listening stations in 2003 was based on our expectation that coho salmon would spawn primarily in Rock Creek; two stations were added in 2004 in response to movement patterns of the fish in 2003. Listening stations were operational each season until it was clear that all movement had ceased.

Mobile tracking by automobile, foot, and raft helped locate tagged fish between listening stations at two different spatial scales: river sections and exact GPS coordinates. Tag positions from both scales were incorporated into a single analysis of movement, as exact locations were nested within river sections. Automobile surveys described the general distribution of tagged salmon above the dam by assigning fish to discrete river sections bounded by predetermined GPS waypoints. The 27 such surveys in 2003 and 25 in 2004 covered an average of 12.8 km (2003) and 11.8 km (2004) of mainstem and tributary habitat. A handheld GPS unit (GPS 76, Garmin, Inc.) defined the way points, and fish were recorded in one of 44 river sections (average length = 554 m, standard deviation (SD) = 190 m).

Foot surveys augmented the automobile tracking to determine precise tag locations or directly observe salmon, but high turbidity, discharge, and the cryptic coloration of spawning coho salmon made direct observations difficult. Consequently, raft surveys (two in 2003, eight in 2004, and two in 2005, including one survey below Landsburg Diversion Dam in 2003 and 2004) were conducted for radio tracking, observations of redds and salmon, and carcass sampling. Direct observations and transmitter detections were recorded as GPS locations for integration with the other types of positional data. Further details on telemetry survey methods can be found in Anderson (2006).

Telemetry data were used to identify specific behaviors and to quantify the overall movement history of individual salmon. The behaviors considered were movement into the tributaries and downstream passage over Landsburg Diversion Dam (2004 only). To relate these two types of movement and Cedar River discharge, we used a GLS procedure identical to that described for counts of upriver movement through the fish ladder. To quantify each individual's movements, the telemetry detections from all sources were organized chronologically and the movements between detections were estimated using ArcGIS version 9.1 (ESRI, Inc., Redlands, California; see Anderson (2006) for further details) to provide minimum estimates of total distance traveled. The last downstream movement was not included because this was assumed to be postmortem drift rather than active movement. Ordinary least squares regression was used to explore the relationship between individual movement and a variety of predictor variables, with separate analyses for males and females. More movement was detected in 2004 than in 2003 because there were more listening stations (six vs. four) and more frequent mobile surveys. To facilitate comparisons between years we standardized the data using a formula analogous to the Z statistic (Zar 1999):

$$(1) \qquad \text{standardized movement } (x) = \frac{x_i - \bar{x}}{\sigma}$$

where x_i is a movement value for fish i, $\bar{x}$ is the average movement within a year, and σ is the standard deviation of movement within a year. Two types of predictor variables were used to explain movement: phenotypic characters (fish length and date of passage at the dam) and conditions encountered on the spawning grounds, which included river discharge (females and males) and availability of ripe females (males only). Both discharge and female availability were averaged over each individual's days of active movement. To estimate the number of ripe females available to each male, we used the counts of females passed upstream at the ladder and assumed that they would be ripe for 1 week based on observations of sexual maturity at the dam.

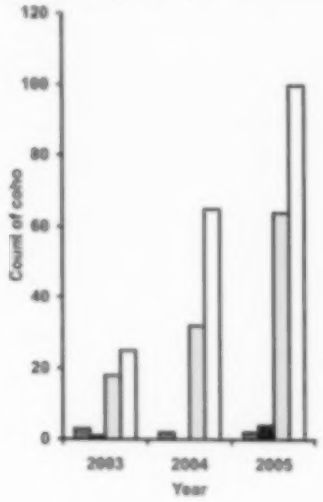
Results

Fish passage

The number of adult coho salmon colonizing the new habitat increased each successive year (Fig. 2), and salmon ascended the fish ladder over a protracted period beginning in early October (Fig. 3). In 2003–2004, 47 coho salmon ascended from 7 October to 21 December (median = 29 November). In 2004–2005, 99 coho salmon migrated upriver through the fish ladder from 11 October to 3 February (median = 13 December). In 2005–2006, 170 coho salmon accessed the new habitat from 9 October to 23 January (median = 20 December). In 2004 and 2005, the migration period extended well past that observed in 2003. Structural modifications made between the 2003 and 2004 seasons at the ladder facility improved late-season capture efficiency, so it is unclear whether this difference was an artifact of sampling or reflected an increase in late season migrants (P. Faulds, Seattle Public Utilities, P.O. Box 34018, Seattle, WA 98124, USA, personal communication). In all three seasons, more males than females entered the new habitat (Fig. 2), but the entry date did not differ between sexes within any

© 2007 NRC Canada

Fig. 2. Number of adult coho salmon (*Oncorhynchus kisutch*) entering the newly accessible habitat by year, sex, and origin (i.e., clipped adipose fin indicated hatchery origin). From left to right within each year: hatchery females (hatched bars), hatchery males (solid bars), naturally spawned females (shaded bars), and naturally spawned males (open bars).



year (*t* tests, $p > 0.05$). Very few coho salmon (3.8% for all years pooled) had clipped adipose fins, indicating that the majority of colonists were naturally produced (Fig. 2). Male size-frequency distributions were unimodal in 2003 and 2004, but in 2005, a second mode below 500 mm fork length likely represented younger, age-2 males, known as jacks (percent of males <500 mm long: 2003, 0.0%; 2004, 1.5%; 2005, 24.0%). Female size-frequency distributions appeared unimodal in all three seasons.

Generalized least squares models revealed that both Cedar River discharge and change in discharge were positively associated with counts at the ladder in 2 of 3 years (discharge: 2003, $p = 0.007$, $r^2 = 0.09$; 2004, $p < 0.0001$, $r^2 = 0.14$; 2005, $p = 0.6$; change in discharge from the previous day: 2003, $p = 0.0001$, $r^2 = 0.16$; 2004, $p = 0.03$, $r^2 = 0.04$; 2005, $p = 0.4$; Fig. 3). Average discharge differed between the years (one-way analysis of variance, ANOVA, $F_{2,338} = 8.0$, $p = 0.0004$), and Tukey tests indicated that 2005 discharges (average = 26.2 $m^3 \cdot s^{-1}$) were higher than those in 2003 (21.0 $m^3 \cdot s^{-1}$, $p = 0.01$) and 2004 (20.0 $m^3 \cdot s^{-1}$, $p = 0.0005$), but discharges in 2003 and 2004 did not differ from each other ($p = 0.8$). Although colder temperatures and higher flows tended to occur toward the end of each season, discharge and temperature were unrelated in a time series regression (GLS, $p > 0.05$). Temperature was not consistently related to upriver movement. In 2003, ladder counts were positively correlated with temperature change from the previous day ($p = 0.002$, $r^2 = 0.11$) but there were no other significant relationships between ladder counts and either temperature or temperature change ($p > 0.05$). Finally, there was little difference in mean temperature between years during the period of coho salmon movement (2003 = 8.3 °C, 2004 = 8.3 °C, 2005 = 7.6 °C; one-way ANOVA, $F_{2,336} = 2.7$, $p = 0.07$).

Spatial distribution patterns

In 2003, 14 females and 23 males received radio tags. In 2004, four females and 45 males received transmitters. In each year, more males were tagged than females, in part because males were more numerous (Fig. 2) and in part because females carried a limited number of gametes and we were concerned that tagging many females may have reduced population productivity through gametes lost via tagging stress. Tagging approximated the overall arrival distribution in 2003, but in 2004, 22 of the 50 fish that did not receive transmitters passed the dam after the final transmitter was implanted, so the tracking data did not represent the latest fish. Nevertheless, the median tagging date matched the median date of all coho salmon passage in both seasons.

Most tagged coho salmon tended to remain in the lower reaches of the Cedar River and few reached Taylor Creek, 12.5 km above the dam (Table 1). In each season, only one individual was detected at the middle Rock Creek site (MR), 2.5 km upstream from the confluence with the Cedar River, and no salmon were detected at the upper Rock Creek (UR) site (Table 1). However, there were many temporary entrances into Rock Creek and two into Taylor Creek (Table 2). Tributary entrances were positively related to discharge in the Cedar River in both years (i.e., entrances were more common when discharge was high; GLS: 2003, $p = 0.003$, $r^2 = 0.10$; 2004, $p = 0.004$, $r^2 = 0.11$). Only two of the 61 entrances were permanent; all others were temporary, and all but four of those were <24 h duration. A higher proportion of males entered tributaries than females (Table 2), and males more often made multiple tributary entrances (Table 2).

All coho salmon redds identified above the dam (5 in 2003, 18 in 2004, and 12 in 2005) were in the Cedar River (almost all in the lower reaches of the newly accessible section), and no redds were in the tributaries. All redds observed in 2003 were <7.0 km from the dam (median = 4.7 km, range = 1.2–6.5 km), as were all but two seen in 2004 (median = 4.8 km, range = 0.9–13.5 km). No radio tracking was conducted in 2005, but surveys of the Cedar River indicated that redds tended to be farther upriver; 50% were >7.0 km above the dam (median = 8.1 km, range = 3.9–12.8 km). Certain areas perennially attracted coho salmon; notably, 31% of all identified redds were in a 700 m section beginning 4.7 km above the dam. Most redds were along the margins of the river, often within 5 m of the bank, and the wetted channel width at spawning sites averaged 36.5 m (SD = 6.5 m, $n = 30$).

Almost half (49%) of the males but fewer females (17%) with transmitters moved below the dam after spending some time upriver (Table 3), as indicated in the 2003 season by a raft survey on 9 March and in 2004 by the listening station at the dam. The 23 fish that moved below the dam in 2004 spent 2.6 h to 37 days (mean = 12.3 days, SD = 9.7 days) in the ladder facility and area above the dam after tagging, and there was no relationship between downstream movement past the dam and Cedar River discharge or discharge change from the previous day (GLS, $p > 0.05$).

Total movement of radio-tagged coho salmon

Patterns of movement within the new habitat differed dramatically between the sexes, as illustrated by the detailed

© 2007 NRC Canada

Fig. 3. Daily counts of adult coho salmon (*Oncorhynchus kisutch*) moving upriver through the fish passage facility with Cedar River mean daily discharge for (*a*) 2003, (*b*) 2004, and (*c*) 2005 seasons. Note the different scales on the *y* axes.

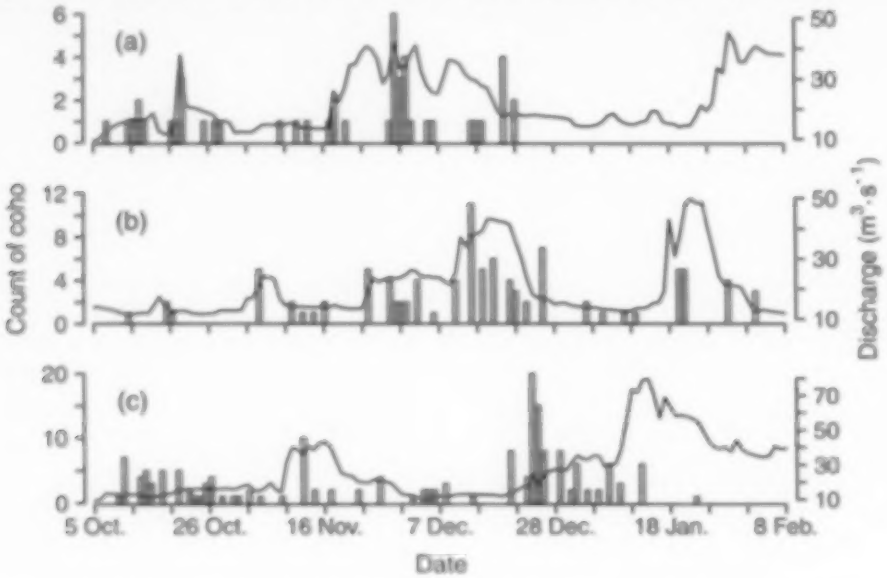


Table 1. Detection of radio-tagged coho salmon (*Oncorhynchus kisutch*) at listening stations, presented as the number of individual fish (*n*) and the proportion of all tagged salmon within that season.

		n		Proportion (%)		Coverage (%)	
Station	Distance (km)	2003	2004	2003	2004	2003	2004
Landsburg (LD)	0.0	—	49	—	100	—	100
Lower Rock (LR)	3.4	30	40	81	82	100	100
Middle mainstem (MM)	6.9	—	36	—	73	—	100
Taylor (TA)	12.5	5	10	14	20	97	100
Middle Rock (MR)	5.9	1	1	3	2	98	95
Upper Rock (UR)	8.0	0	0	0	0	81	50

Note: Distance refers to the longitudinal distance in river kilometres from the tag site at Landsburg Diversion Dam to the position of the listening station. Coverage represents the proportion of time during each season that the listening station was operational.

Table 2. Entrances by radio-tagged coho salmon (*Oncorhynchus kisutch*) into tributaries of the Cedar River.

	2003		2004	
	Males	Females	Males	Females
Number of individual salmon	12	3	17	1
Proportion of total tagged (%)	52	21	38	25
Number of total entrances	34	3	23	1
Mean entrances per individual	2.8	1.0	1.4	1.0
Proportion temporary (%)	97	67	100	100
Duration mean (h)	5.1	2.7	7.9	2.2
Duration standard deviation (h)	6.0	0.1	11.1	—
Duration maximum (h)	24.8	2.8	38.6	2.2

Note: With only two exceptions, these represent entrances into Rock Creek. The two exceptions occurred in 2003, when two different males each entered Taylor Creek once.

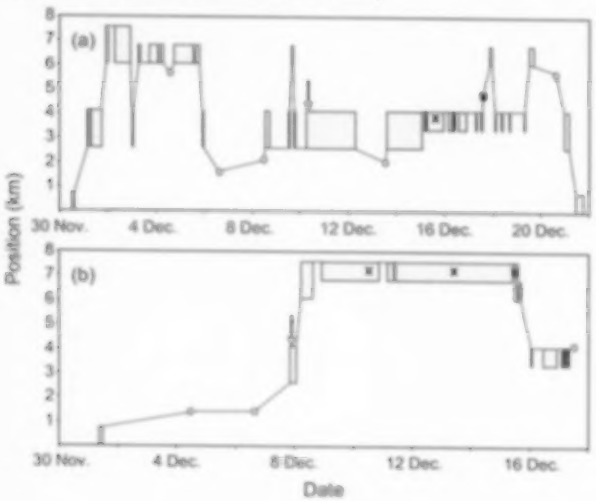
movement histories of one representative male and female (Fig. 4). The male's total distance moved (45.8 km) was very similar to that of the male population median in that season (45.0 km) even though he was never detected more than 7.6 km upriver from the dam. Instead, he moved frequently within that stretch of river until exiting below the

© 2007 NRC Canada

Table 3. Final locations of radio-tagged coho salmon (*Oncorhynchus kisutch*) with respect to Landsburg Diversion Dam.

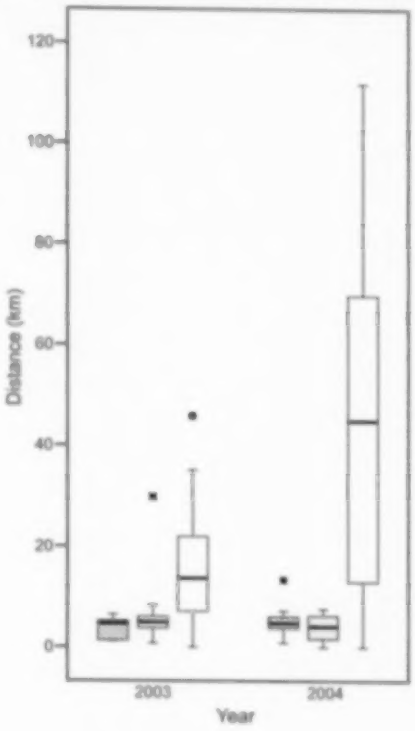
Sex	Year	Above dam	Below dam	Unknown
Males	2003	10	10	3
	2004	23	22	0
Females	2003	12	2	0
	2004	3	1	0

Fig. 4. Detailed movement history of typical (*a*) male and (*b*) female coho salmon (*Oncorhynchus kisutch*). Date and time are along the *x* axis, and longitudinal position within the Cedar River is on the *y* axis. The boxes are listening station detections, where the width is proportional to the duration and the height is proportional to the estimated detection range. Boxes of half height are detections where an individual remained exclusively above or below the station itself. Mobile detections are represented by circles. Arrows are temporary tributary entrances, and "×" denotes a direct visual observation of the animal on a redd. The male was observed on two different redds, whereas the female was observed on a single redd multiple times. Movement distance for the male was summed up to the final detection at the MM station on 19 December. This individual passed downriver past Landsburg Diversion Dam on 21 December. The female in the bottom panel was detected repeatedly in the same location as shown by the final mobile detection plotted on 17 December before the carcass was recovered at this same location on 7 January.



dam 21 days after entry. The female moved upriver more slowly, briefly explored Rock Creek, and established her redd site 7.2 km above the dam approximately 8 days after tagging. She never undertook the extensive back and forth movements characteristic of most males (total distance moved = 7.6 km; 2004 female population median = 4.1 km) and remained in the vicinity of her redd site until the final downriver movement that was assumed to be carcass drift 14 days after ascending the ladder. Males (average minimum estimate = 34.8 km, $n = 66$) moved much more than females (5.8 km, $n = 18$; Fig. 5). Estimates of minimum movement

Fig. 5. Coho salmon (*Oncorhynchus kisutch*) spawning site distances and total movement of radio-tagged colonists within the newly accessible habitat. From left to right within each year are redd distances (hatched boxes, $n = 5$ in 2003, $n = 18$ in 2004), female movement (shaded boxes, $n = 14$ in 2003, $n = 4$ in 2004), and male movement (open boxes, $n = 21$ in 2003, $n = 45$ in 2004). In each plot, the thick middle bar is the median, the boxes represent the interquartile range, and the remainder of the data is contained in the whiskers and outliers points (solid squares). Redd distances are measured from Landsburg Diversion Dam in river kilometres. Comparing movement between seasons is not valid because of the addition of two listening stations and increased frequency of mobile telemetry surveys in 2004.

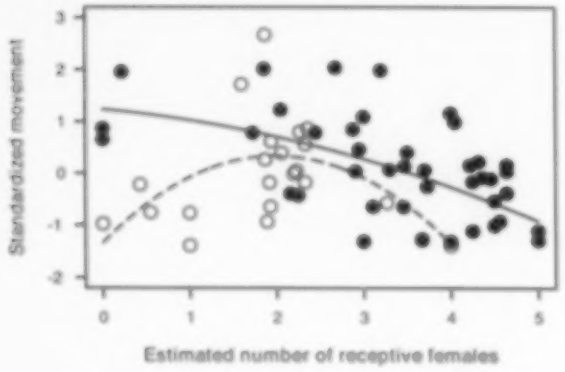


were higher in 2004, primarily because the two additional telemetry stations increased our power to detect movement, but males moved more than females in both 2003 (Mann–Whitney $U = 59.0$, $p = 0.003$) and 2004 (Mann–Whitney $U = 28.5$, $p = 0.03$). Median total movement of females within the new habitat was approximately equal to the median of the upriver distance of identified redds (Fig. 5).

Female movement was not significantly related to migration date, body length, or Cedar River discharge (single predictor variable ordinary least squares regression, all $p > 0.05$). Similarly, male movement was unrelated to migration date, body length, and discharge (all $p > 0.05$). However, male movement was negatively related to the estimated number of ripe females within the new habitat ($p = 0.008$, $r^2 = 0.10$; Fig. 6); males moved more when females were scarce. To further explore the relationship between estimated count of ripe females and male movement, we compared linear and quadratic ($y = ax^2 + bx + c$) models. The quadratic model

© 2007 NRC Canada

Fig. 6. Relationship between estimated number of receptive females and total movement by individual male coho salmon (*Oncorhynchus kisutch*) in 2003 (open circles, $n = 21$) and 2004 (solid circles, $n = 45$). Quadratic models of total movement are shown as a broken line for 2003 ($p = 0.04$, $r^2 = 0.29$) and a solid line for 2004 ($p = 0.004$, $r^2 = 0.31$). On the y axis, zero represents average movement within each year, and the values are the number of standard deviations away from the annual mean.



better fit the 2003 data ($n = 21$; linear, $p = 0.7$; quadratic, $p = 0.04$, $r^2 = 0.29$), but the models performed equally well for the 2004 data ($n = 45$; linear, $p < 0.0001$, $r^2 = 0.30$; quadratic, $p = 0.0004$, $r^2 = 0.31$). Tagged males had access to more ripe females in 2004 (average = 3.4) compared with 2003 (average = 1.9). The linear model performed well when there were >3 females (densities this high were rare in 2003 but more common in 2004) but poorly when females were scarce. The quadratic model performed well in 2003 because it captured a dome-shaped relationship (Fig. 6); male movement was greatest at intermediate levels of female availability. Thus there was strong evidence from both years that males seldom moved when females were numerous, and weak evidence from 2003 that male movement was also reduced when females were scarce (Fig. 6).

Discussion

Dam passage

The first objective was to document the basic ecological attributes of coho salmon entering newly accessible habitat. Colonization began in the very first season when Landsburg Diversion Dam was passable; 47 coho salmon moved upriver in fall 2003, increasing to 99 in 2004, and 170 in 2005. Obviously these original colonists did not return to their natal site, as the habitat extension afforded by the passage facilities permitted range expansion by salmon spawned in the Cedar River below the dam and straying from elsewhere. Although the origin of the colonists is unknown, it seems reasonable to assume that many, if not most, were produced in the 35 river kilometres of the Cedar River or its tributaries below the dam. Very few had clipped adipose fins indicating hatchery origin, and both hatcheries in the Lake Washington basin (University of Washington and Issaquah Creek) clipped at least 93% of their juvenile coho salmon in the appropriate release years (Regional Mark Processing Center database, www.rmpc.org). Thus the vast majority of colonists were naturally spawned. The counts of coho salmon ascending the ladder increased in successive years, but the lack of comprehensive abundance data on coho salmon below the dam makes it impossible to determine whether this reflects higher rates of colonization or merely increases in coho salmon throughout the system.

Basic population demographics revealed three interesting observations. First, as a result of low abundances and protracted timing, instantaneous spawning densities were extremely low (average daily estimate of ripe females above dam = 1.8 in 2003, 1.8 in 2004, and 4.5 in 2005). Second, in all three seasons, more males entered the new habitat than females. This may reflect, in part, the generally higher survival of male coho salmon at sea and overabundance relative to females in many populations (Holtby and Healey 1990; Spidle et al. 1998). In addition, males may have been more likely to enter the fish passage facility because of their elevated movement rates compared with that of females. Third, we observed a large proportion of small males, presumably age-2 jacks, in 2005. The sudden abundance of this age class in 2005 after 2 years of absence raises the possibility that these fish were the offspring of the original 2003 brood year, homing to natal sites above the dam. However, the proportion of jacks within a population varies between years (Vøllestad et al. 2004), so the abundance of jacks in 2005 may have resulted from processes unrelated to homing such as smolt size or overall year-class strength.

Passage results were also used to evaluate the proximate environmental stimuli promoting upriver movement through the ladder. The idea that higher stream flows induce upriver movement by adult salmon has been discussed for decades (Davidson et al. 1943; Shapovalov and Taft 1954; Banks 1969), though fish may respond to changes in flow rather than the absolute value (Trépanier et al. 1996). In two years, both absolute discharge and change from the previous day were positively related to ladder counts, but no discharge effect was observed in 2005, likely because many of the high flow days occurred toward the end of the season when few fish were left to enter. Similarly, Erkinaro et al. (1999) noted that point-checking data, such as the ladder counts used in this study, will tend to underestimate the effects of environmental stimuli if there are no fish available downstream of the counting location. In contrast to discharge, temperature showed no clear effect on passage.

Spawning sites

The second objective was to assess reach-scale patterns in spawning site selection and evaluate the hypothesis that coho colonists would spawn primarily in tributaries. Coho salmon spawn in small streams elsewhere in the Lake Washington basin and in tributaries of the Cedar River below the dam, and this species is typically found in small streams of intermediate gradient (Sandercock 1991), so we expected extensive use of Rock Creek for spawning. We cannot rule out the possibility that untagged coho salmon spawned in tributaries but we found no evidence of tagged fish spawning in tributaries despite continuous monitoring of Rock Creek and Taylor Creek via telemetry listening stations and periodic mobile tracking surveys of all four tributaries. Listening station data showed that 38% of tagged salmon entered Rock

© 2007 NRC Canada

Creek for temporary visits, but these trips were of insufficient duration for spawning. These temporary entrances indicated that the lack of tributary spawning was not because salmon were unable to locate tributaries; these areas were simply not selected. Furthermore, snorkel surveys indicated that juvenile coho salmon were absent from Rock Creek for much of the spring and early summer but apparently entered from the mainstem later in the growing season (Anderson 2006). Unfortunately, the lack of information on the relative frequency of tributary versus mainstem spawning below the dam prevents quantitative comparisons to breeding habitat preferences elsewhere in the watershed.

There are at least two possible explanations for why coho salmon did not spawn in the tributaries, as seems to be typical for the species (Sandercock 1991). First, the Cedar River may represent higher-quality adult spawning habitat and mainstem redd locations resulted from behavioral choice. The temporary exploration of Rock Creek by 38% of tagged salmon in both telemetry seasons is consistent with this hypothesis, as fish accessed the tributary habitat but did not spawn there. The second possibility is that coho salmon colonists did not imprint on the odors of the tributaries and thus were more reluctant to settle there. Under this scenario, the Cedar River below the dam was the source of colonists. Within the range of odors presented by the new habitat, Cedar River odors were more similar to those of natal sites below the dam than those of the tributaries, especially if the fish had been spawned in the Cedar River itself rather than in the tributaries below the dam. For example, the concentration of total organic carbon in Rock Creek (mean ± 1 SD = 2.9 ± 1.4 mg·L^{-1}) was more than four times greater than in the Cedar River (measured just upriver from the Rock Creek confluence: 0.7 ± 0.2 mg·L^{-1}; P. Kiffney, National Marine Fisheries Service, 2725 Montlake Blvd East, Seattle, WA 98112, USA, unpublished data). Perhaps the colonists identified differences in organic carbon or other chemical signals and chose spawning locations based on olfaction despite the absence of authentic homing cues, overcoming a general attraction to tributaries for breeding.

Movements of individual salmon

Our third objective was to quantify individual movement within the new habitat and evaluate the prediction that movement would be inversely correlated with breeding opportunities. Movements of radio-tagged salmon were highly variable, as some individuals moved over 100 km within the habitat, whereas others spent only hours above the dam before returning downstream. The most striking pattern was that males moved significantly more than females. Sex-based movement differences have been previously observed (McCubbing et al. 1998; Hutchings and Gerber 2002) and likely can be explained by differing sources of breeding competition. Female salmon compete with each other for access to spawning territories (Schroder 1981; van den Berghe and Gross 1989; Foote 1990). Densities of female salmon in our study were very low and potential spawning sites were abundant, so females had little need for widespread movement. Males, on the other hand, compete with each other for access to ripe females (Schroder 1981; Foote 1990; Quinn et al. 1996). With few potential mates, far-ranging movement may have been needed to find a receptive female. Competitively inferior males may have been displaced from available mates, and then moved to find a female, perhaps unsuccessfully, for which they could compete (Healey and Prince 1998). Furthermore, males often depart females in search of other breeding opportunities immediately after a spawning event, whereas females remain to guard the nest site (Foote 1990). Males moved more when receptive females were scarce, further supporting the prediction that breeding opportunity and movement would be inversely related. There appeared to be a threshold effect, whereby male movement was extremely variable and unrelated to female availability until there were at least three females to court. Beyond this threshold, female abundance had a clear inverse relationship with male movement.

The total movement values we report for male colonists vastly exceed those of other studies recording salmon movement on the spawning grounds, and we propose two potential explanations. First, the colonists in this study may have been a nonrandom subset of the source population, predisposed to extensive movements because they were competitively inferior. Healey and Prince (1998) observed that satellite males moved more widely than alpha males. Perhaps the colonists in the present study entered the new habitat because they were excluded from breeding opportunities below the dam. Second, observed spawning densities in the Cedar River were much lower than in previous movement studies of >1000 fish·km^{-1} for stream-spawning sockeye (Stewart et al. 2004; Rich et al. 2006), >0.5 fish·m^{-2} for beach-spawning sockeye (Hendry et al. (1995), and approximately 150 fish·km^{-1} for coho spawning in a small creek (Healey and Prince 1998). We observed an effect of density, and male movement might have been further reduced at even higher female densities.

Adult salmon movements can be divided into different phases: (*i*) migratory, a period of comparatively rapid upriver movement, (*ii*) search, a phase of erratic upriver and downriver movements (Økland et al. 2001; Connor and Garcia 2006), and (*iii*) spawning movements. The movements observed here exceeded those previously reported for spawning behavior, but they were quite similar to those reported for search behavior. Migrating salmon move much more rapidly (10–50 km·day^{-1}; Keefer et al. 2004) than fish on spawning grounds (0–1 km·day^{-1}; Healey and Prince 1998). Many of the far-ranging movements by males reported here (average minimum estimate = 34.8 km) were repeated upriver and downriver excursions and were similar to search distances previously reported for Chinook (wild male average = 39 km; Connor and Garcia 2006) and Atlantic salmon (*Salmo salar*; both sexes, 1992 average = 30.9 km, 1993 average = 22.8 km; Økland et al. 2001).

Body size, entry date, and discharge were not significantly correlated to the total movement of individual fish, although many salmon were present during periods of increased discharge, particularly in 2004. Consequently, males may have moved less at higher flows because more females were available to court. However, movement into tributaries was positively related to Cedar River discharge in both telemetry seasons. This movement was predominantly temporary, and fish may have been escaping high flows encountered in the

© 2007 NRC Canada

mainstem, exploring potential breeding habitat, or searching for mates. Thus the results for discharge were mixed; salmon moved upriver through the ladder and into tributaries during periods of increased flow but discharge was not related to the total movements of radio-tagged individuals. Discharge likely provided a stimulus for upriver movement before fish reached sexual maturity. Once reproductive behavior dominated movements, discharge was no longer an important factor, consistent with the idea that discharge and sexual maturity interact to affect adult salmon movement (Davidson et al. 1943).

In both seasons, roughly half of the tagged males, and a lower proportion of tagged females, moved below Landsburg Diversion Dam after spending some time upriver. It is unlikely that this is purely the result of carcass drift, as carcasses tended to settle in a large, deep pool immediately upriver from the dam. Moreover, if downstream movement was the result of carcass drift, it would have been positively related to discharge and occurred in equal proportions for both sexes, but neither result was observed. The frequency of downstream passage by males was likely related to their higher overall movement rate, as they may have moved downriver in search of mates. Many Chinook salmon passing downriver over Columbia River dams later entered spawning grounds or hatcheries, suggesting that salmon may retreat if they "overshoot" their natal site (Boggs et al. 2004). Interestingly, the lone radio-tagged female in 2004 that moved below the dam spent approximately 6 h above the dam before retreating downriver and was observed on a redd, courted by two males 3.1 km below Landsburg Dam, 5 days after tagging.

Colonization, homing, and conservation

In this initial phase of population expansion, all colonists were strays; they could not be homing precisely to their natal site because this area had not been accessible for the last 100 years. Salmon accessing the new habitat either failed to home accurately or strayed by choice. Salmon use olfactory cues to identify their natal site (Dittman and Quinn 1996). A failure to home could have occurred if there were no odor differences between the Cedar River above and below the dam, or if fish could not detect the differences. Previous studies have demonstrated that salmon are capable of detecting differences in olfactory cues at the reach scale within a single creek or river (Wagner 1969; Stewart et al. 2004; Quinn et al. 2006). Local geology will likely have a strong influence on site-specific odors, and thus the spatial scale at which detectable differences occur will vary substantially between watersheds. Alternatively, the odors of conspecifics could have played a role in the colonization process. In each successive year of coho salmon passage, increased adult counts followed increased densities of juvenile coho salmon observed the previous summer (Anderson 2006). This observation is consistent with the pheromone hypothesis advanced by Nordeng (1971; 1977), as adult colonists may have received an olfactory cue from the previous year's offspring. Experimental studies indicated that adult coho salmon were attracted to the odors of juvenile coho salmon (Quinn et al. 1983), but it is unclear to what extent these odors affected movements in the Cedar River. It is also possible that changes in coho salmon abundance within the entire system caused the increasing numbers of colonists.

In addition to the physiological explanations, salmon may have strayed into the new habitat by choice. Female salmon colonists had virtually no competition for breeding space (e.g., van den Berghe and Gross 1989), so they could devote their energy almost entirely to nest site selection and preparation rather than territory acquisition and defense. In addition, they encountered a biologically and physically well-developed stream community, with abundant pools, instream woody debris, and prey base for offspring. This high-quality habitat contrasts with the young streams made available by retreating glaciers (Milner and Bailey 1989; Milner et al. 2000), and it would be interesting to compare movements of salmon in these systems. There likely exists a trade-off between precise homing and spawning habitat selection whereby salmon establish nests away from their natal site if local conditions elsewhere are better (Stewart et al. 2004). In this case, female colonists may have detected differences from odors imprinted elsewhere, but settled above the dam because of the low competition and high habitat quality.

In conclusion, we have documented the initial process of population expansion whereby adult salmon colonized an area from which they had been extirpated for over 100 years. The number of salmon that entered the new habitat via the dam passage facility, as well as the widespread movements of some salmon above the dam, suggests that exploration is an innate component of salmon breeding behavior. The more restricted movements of females compared with males, in addition to their smaller numbers of gametes, suggests that female dispersal will likely limit the numerical and spatial expansion of the population in the first few generations. Clearly, dispersal ability benefits salmon when new habitat becomes available, but Pacific salmon also possess other traits identified by Safriel and Ritte (1983) as common to good colonizers, including rapid growth, early reproductive maturity, and large reproductive effort (Quinn 2005). The establishment of a self-sustaining population following colonization will also depend on the reproductive success of salmon after arriving in the new habitat. Thus, in the Cedar River, the survival of offspring rearing in a community of competitive and predatory resident fish will likely determine the population's future trajectory.

The ability to exploit new habitat implies that dam circumvention projects have great potential to benefit salmon populations. Freshwater habitat loss threatens salmon throughout much of their range, and dams have dramatically reduced access to spawning and rearing areas (Nehlsen et al. 1991). Agencies tasked with aiding the recovery of salmon may take a variety of approaches. Although artificial supplementation offers one option to seed unoccupied habitats (Young 1999), genetic risks of translocation may outweigh potential benefits (Utter 2001; Eldridge and Naish 2007). Alternatively, stream restoration can recuperate natural fluvial and biological processes in lotic ecosystems to the benefit of salmon. Despite substantial funding for such projects, limited evaluation of outcomes has restricted our ability to learn from mistakes and achievements (Bernhardt et al. 2005). In planning stream restoration activities, management should prioritize reconnection of isolated high-quality habitats (Roni et al. 2002). Our results demonstrate that salmon

© 2007 NRC Canada

are apt to disperse into new areas, and continued documentation of natural recolonization rates and processes will help managers identify migration barriers suitable for removal or circumvention. Restoring access to lost habitats will provide opportunities for salmon to explore, colonize, and recover.

Acknowledgements

We thank Seattle Public Utilities, in particular P. Faulds and J. McDowell for sampling adult salmon at Landsburg Diversion Dam and B. Bachen for continued support of collaborative research. Many individuals helped with fieldwork, but K. Burton, P. Westley, and A. Kingham deserve special recognition. S. Carlson, D. Chapin, R. Little, D. Paige, and two anonymous reviewers provided valuable comments to earlier versions of this manuscript. This work was funded in part by a grant from Washington Sea Grant, University of Washington, pursuant to National Oceanic and Atmospheric Administration Award No. NA04OAR4170032, Project No. R/F-148. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies. Seattle Public Utilities, the Washington Fly Fishing Club, and the H. Mason Keeler Endowment provided additional funding for this work.

References

Anderson, J.H. 2006. Colonization of newly accessible habitat by coho salmon (*Oncorhynchus kisutch*). M.S. thesis, School of Aquatic and Fishery Sciences, University of Washington, Seattle, Wash.

Banks, J.W. 1969. A review of the literature on the upstream migration of adult salmonids. J. Fish Biol. **1**: 85–136.

Berman, C.H., and Quinn, T.P. 1991. Behavioural thermoregulation and homing by spring chinook salmon, *Oncorhynchus tshawytscha* (Walbaum), in the Yakima River. J. Fish Biol. **39**: 301–312.

Bernhardt, E.S., Palmer, M.A., Allan, J.D., Alexander, G., Barnas, K., Brooks, S., Carr, J., Clayton, S., Dahm, C., Follstad-Shah, J., Galat, D., Gloss, S., Goodwin, P., Hart, D., Hassett, B., Jenkinson, R., Katz, S., Kondolf, G.M., Lake, P.S., Lave, R., Meyer, J.L., O'Donnell, T.K., Pagano, L., Powell, B., and Sudduth, E. 2005. Synthesizing U.S. river restoration efforts. Science (Washington, D.C.), **238**: 636–637.

Boggs, C.T., Keefer, M.L., Peery, C.A., Bjornn, T.C., and Stuehrenberg, L.C. 2004. Fallback, reascension and adjusted fishway escapement estimates for adult Chinook salmon and steelhead at Columbia and Snake River dams. Trans. Am. Fish. Soc. **133**: 932–949.

Ciancio, J.E., Pascual, M.A., Lancelotti, J., Riva Rossi, C.M., and Botto, F. 2005. Natural colonization and establishment of a chinook salmon, *Oncorhynchus tshawytscha*, population in the Santa Cruz River, an Atlantic basin of Patagonia. Environ. Biol. Fishes, **74**: 219–227.

Connor, W.P., and Garcia, A.P. 2006. Prespawning movement of wild and hatchery fall Chinook salmon adults in the Snake River. Trans. Am. Fish. Soc. **135**: 297–305.

Cury, P. 1994. Obstinate nature: an ecology of individuals. Thoughts on reproductive behavior and biodiversity. Can. J. Fish. Aquat. Sci. **51**: 1664–1673.

Davidson, F.A., Vaughan, E., and Hutchinson, S.J. 1943. Factors influencing the upstream migration of the pink salmon (*Oncorhynchus gorbuscha*). Ecology, **24**: 149–168.

Dittman, A.H., and Quinn, T.P. 1996. Homing in Pacific salmon: mechanisms and ecological basis. J. Exp. Biol. **199**: 83–91.

Eldridge, W.H., and Naish, K.A. 2007. Long-term effects of translocation and release numbers on fine-scale population structure among coho salmon (*Oncorhynchus kisutch*). Mol. Ecol. **16**: 2407–2421.

Erkinaro, J., Økland, F., Moen, K., Niemelä, E., and Rahiala, M. 1999. Return migration of Atlantic salmon in the River Tana: the role of environmental factors. J. Fish Biol. **55**: 506–516.

Foote, C.J. 1990. An experimental comparison of male and female spawning territoriality in a Pacific salmon. Behaviour, **115**: 283–314.

Healey, M.C., and Prince, A. 1998. Alternative tactics in the breeding behaviour of male coho salmon. Behaviour, **135**: 1099–1124.

Hendry, A.P., Leonetti, F.E., and Quinn, T.P. 1995. Spatial and temporal isolating mechanisms: the formation of discrete breeding aggregations of sockeye salmon (*Oncorhynchus nerka*). Can. J. Zool. **73**: 339–352.

Holtby, L.B., and Healey, M.C. 1990. Sex-specific life history tactics and risk-taking in coho salmon. Ecology, **71**: 678–690.

Hutchings, J.A., and Gerber, L. 2002. Sex-biased dispersal in a salmonid fish. Proc. R. Soc. Biol. Sci. Ser. B, **269**: 2487–2493.

Keefer, M.L., Peery, C.A., Jepson, M.A., and Stuehrenberg, L.C. 2004. Upstream migration rates of radio-tagged adult chinook salmon in riverine habitats of the Columbia River basin. J. Fish Biol. **65**: 1126–1141.

Keefer, M.L., Peery, C.A., and Caudill, C.C. 2006. Long-distance downstream movements by homing adult chinook salmon. J. Fish Biol. **68**: 944–950.

Kiffney, P., Volk, C., Hall, J., and Eberhart, C. 2002. Community and ecosystem attributes of the Cedar River watershed above Landsburg Diversion before the arrival of Pacific salmon. Technical report submitted to Seattle Public Utilities, National Marine Fisheries Service, Northwest Fisheries Science Center Watershed Program, Seattle, Wash.

Kwain, W. 1987. Biology of pink salmon in the North American Great Lakes. Am. Fish. Soc. Symp. **1**: 57–65.

LePage, C., and Cury, P. 1997. Population viability and spatial fish reproductive strategies in constant and changing environments: an individual-based modelling approach. Can. J. Fish. Aquat. Sci. **54**: 2235–2246.

McCubbing, D.J.F., Bayliss, B.D., and Locke, V.M. 1998. Spawning migration of radio-tagged landlocked Arctic charr, *Salvelinus alpinus* L., in Ennerdale Lake, the English Lake District. Hydrobiologia, **371/372**: 173–180.

McPhail, J.D., and Lindsey, C.C. 1986. Zoogeography of the freshwater fishes of Cascadia (the Columbia system and rivers north to the Stikine). *In* The zoogeography of North American freshwater fishes. *Edited by* C.H. Hocutt and E.O. Wiley. John Wiley & Sons, New York. pp. 615–637.

Milner, A.M., and Bailey, R.G. 1989. Salmonid colonization of new streams in Glacier Bay National Park, Alaska. Aquacult. Fish. Manag. **20**: 179–192.

Milner, A.M., Knudsen, E.E., Soiseth, C., Robertson, A.L., Schell, D., Phillips, I.T., and Magnusson, K. 2000. Colonization and development of stream communities across a 200-year gradient in Glacier Bay National Park, Alaska, U.S.A. Can. J. Fish. Aquat. Sci. **57**: 2319–2335.

Montgomery, D.R., Beamer, E.M., Pess, G.R., and Quinn, T.P. 1999. Channel type and salmonid spawning distribution and abundance. Can. J. Fish. Aquat. Sci. **56**: 377–387.

Nagelkerke, N.J.D. 1991. A note on a general definition of the coefficient of determination. Biometrika, **78**: 691–692.

© 2007 NRC Canada

Nehlsen, W., Williams, J.E., and Lichatowich, J.A. 1991. Pacific salmon at the crossroads: stocks at risk from California, Oregon, Idaho, and Washington. Fisheries, **16**: 4–21.

Newell, J.C., and Quinn, T.P. 2005. Behavioral thermoregulation by maturing adult sockeye salmon (*Oncorhynchus nerka*) in a stratified lake prior to spawning. Can. J. Zool. **83**: 1232–1239.

Nordeng, H. 1971. Is the local orientation of anadromous fishes determined by pheromones? Nature (London), **233**: 411–413.

Nordeng, H. 1977. A pheromone hypothesis for homeward migration in anadromous salmonids. Oikos, **28**: 155–159.

Økland, F., Erkinaro, J., Moen, K., Niemelä, E., Fiske, P., McKinley, R.S., and Thorstad, E.B. 2001. Return migration of Atlantic salmon in the River Tana: phases of migratory behaviour. J. Fish Biol. **59**: 862–874.

Ostrom, C.W. 1990. Time series analysis: regression techniques. Sage University Paper Series on Quantitative Applications in the Social Sciences No. 07–009, Sage, Newbury Park, Calif.

Peters, R.J. 1996. An evaluation of habitat enhancement and wild fry supplementation as a means of increasing coho salmon production of the Clearwater River, Washington. Ph.D. dissertation, School of Fisheries, University of Washington, Seattle, Wash.

Quinn, T.P. 1993. A review of homing and straying in wild and hatchery-produced salmon. Fish. Res. **18**: 29–44.

Quinn, T.P. 2005. The behavior and ecology of Pacific salmon and trout. University of Washington Press, Seattle, Wash.

Quinn, T.P., Brannon, E.L., and Whitman, R.P. 1983. Pheromones and the water source preferences of adult coho salmon, *Oncorhynchus kisutch* Walbaum. J. Fish Biol. **22**: 677–684.

Quinn, T.P., Adkison, M.D., and Ward, M.B. 1996. Behavioral tactics of male sockeye salmon (*Oncorhynchus nerka*) under varying operational sex ratios. Ethology, **102**: 304–322.

Quinn, T.P., Kinnison, M.T., and Unwin, M.J. 2001. Evolution of chinook salmon (*Oncorhynchus tshawytscha*) populations in New Zealand: pattern, rate, and process. Genetica, **112–113**: 493–513.

Quinn, T.P., Stewart, I.J., and Boatright, C.P. 2006. Experimental evidence of homing to site of incubation by mature sockeye salmon, *Oncorhynchus nerka*. Anim. Behav. **72**: 941–949.

Rich, H.B., Carlson, S.M., Chasco, B.E., Briggs, K.C., and Quinn, T.P. 2006. Movement of male sockeye salmon, *Oncorhynchus nerka*, on spawning grounds: effects of in-stream residency, density, and body size. Anim. Behav. **71**: 971–981.

Ricker, W.E. 1972. Hereditary and environmental factors affecting certain salmonid populations. *In* The stock concept in Pacific salmon. *Edited by* R.C. Simon and P.A. Larkin. H.R. MacMillan Lectures in Fisheries, Vancouver, B.C. pp. 19–160.

Rieman, B.E., and Dunham, J.B. 2000. Metapopulations and salmonids: a synthesis of life history patterns and empirical observations. Ecol. Freshw. Fish. **9**: 51–64.

Roni, P., Beechie, T.J., Bilby, R.E., Leonetti, F.E., Pollock, M.M., and Pess, G.R. 2002. A review of stream restoration techniques and a hierarchical strategy for prioritizing restoration in Pacific Northwest watersheds. N. Am. J. Fish. Manag. **22**: 1–20.

Rosenfeld, J., Porter, M., and Parkinson, E. 2000. Habitat factors affecting the abundance and distribution of juvenile cutthroat trout (*Oncorhynchus clarki*) and coho salmon (*Oncorhynchus kisutch*). Can. J. Fish. Aquat. Sci. **57**: 766–774.

Safriel, U.N., and Ritte, U. 1983. Universal correlates of colonizing ability. *In* The ecology of animal movement. *Edited by* I.R. Swingland and P.J. Greenwood. Clarendon Press, Oxford, UK. pp. 215–239.

Sandercock, F.K. 1991. Life history of coho salmon (*Oncorhynchus kisutch*). *In* Pacific salmon life histories. *Edited by* C. Groot and L. Margolis. UBC Press, Vancouver, B.C. pp. 395–445.

Scarnecchia, D.L., and Roper, B.B. 2000. Large-scale, differential summer habitat use of three anadromous salmonids in a large river basin in Oregon, USA. Fish. Manag. Ecol. **7**: 197–209.

Schroder, S.L. 1981. The role of sexual selection in determining overall mating patterns and mate choice in chum salmon. Ph.D. dissertation, College of Fisheries, University of Washington, Seattle, Wash.

Seattle Public Utilities 2000. Final Cedar River watershed habitat conservation plan. City of Seattle, Seattle, Wash.

Shapovalov, L., and Taft, A.C. 1954. The life histories of the steelhead rainbow trout (*Salmo gairdneri gairdneri*) and silver salmon (*Oncorhynchus kisutch*) with special reference to Waddell Creek, California, and recommendations regarding their management. Fish Bull. 98, California Department of Fish and Game, Sacramento, Calif.

Spidle, A.P., Quinn, T.P., and Bentzen, P. 1998. Sex-biased marine survival and growth in a population of coho salmon. J. Fish Biol. **52**: 907–915.

Stewart, I.J., Carlson, S.M., Boatright, C.P., Buck, G.B., and Quinn, T.P. 2004. Site fidelity of spawning sockeye salmon (*Oncorhynchus nerka* W.) in the presence and absence of olfactory cues. Ecol. Freshw. Fish. **13**: 104–110.

Taylor, E.B. 1991. A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. Aquaculture, **98**: 185–207.

Trépanier, S., Rodríguez, M.A., and Magnan, P. 1996. Spawning migrations in landlocked Atlantic salmon: time series modelling of river discharge and water temperature effects. J. Fish Biol. **48**: 925–936.

Utter, F. 2001. Patterns of subspecific anthropogenic introgression in two salmonid genera. Rev. Fish Biol. Fish. **10**: 265–279.

van den Berghe, E.P., and Gross, M.R. 1989. Natural selection resulting from female breeding competition in a Pacific salmon (coho: *Oncorhynchus kisutch*). Evolution, **43**: 125–140.

Vøllestad, L.A., Peterson, J., and Quinn, T.P. 2004. Effects of freshwater and marine growth rates on early maturity in male coho and Chinook salmon. Trans. Am. Fish. Soc. **133**: 495–503.

Wagner, H.H. 1969. Effect of stocking location of juvenile steelhead trout, *Salmo gairdnerii*, on adult catch. Trans. Am. Fish. Soc. **98**: 27–34.

Young, K.A. 1999. Managing the decline of Pacific salmon: metapopulation theory and artificial recolonization as ecological mitigation. Can. J. Fish. Aquat. Sci. **56**: 1700–1706.

Zar, J.H. 1999. Biostatistical analysis. Prentice-Hall, Upper Saddle River, N.J.

© 2007 NRC Canada

ERRATUM / ERRATUM

Erratum: Admitting ageing error when fitting growth curves: an example using the von Bertalanffy growth function with random effects

Jason M. Cope and André E. Punt

Ref.: Can. J. Fish. Aquat. Sci. **64**(2): 205–218 (2007).

Equation 1 was incorrectly written as

$$l_i + L_i + \varepsilon_{L,i} \qquad \varepsilon_{L,i} \sim N(0, \sigma^2_{L,i})$$

Equation 1 should read as

$$l_i = L_i + \varepsilon_{L,i} \qquad \varepsilon_{L,i} \sim N(0, \sigma^2_{L,i})$$

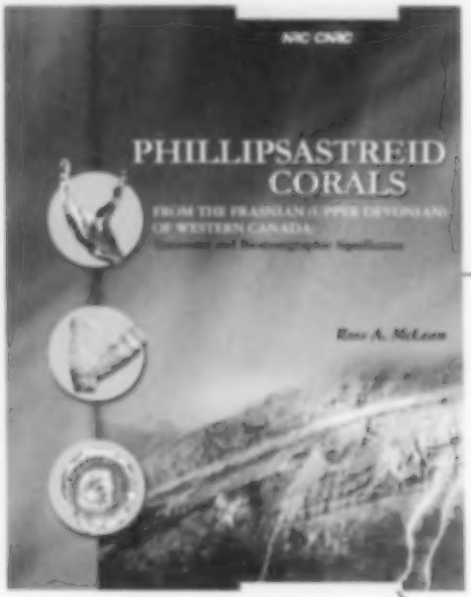
Received 20 July 2007. Accepted 20 July 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 9 August 2007. J19291E

J.M. Cope[1] and A.E. Punt. School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA.

[1]Corresponding author (e-mail: jcope@u.washington.edu).

 doi:10.1139/F07-095 © 2007 NRC Canada

National Research Council Canada | Conseil national de recherches Canada



Phillipsastreid Corals

from the Frasnian (Upper Devonian) of Western Canada: Taxonomy and Biostratigraphic Significance

By / Par
Ross A. McLean

109 Pages 2005
Softcover / Édition brochée

ISBN 0-660-19364-7; NRC 46847
ISBN 0-660-19364-7; CNRC 46847

$49.95 CAN; Other countries: $49.95 US
49,95 $CAN; À l'étranger : 49,95 $US

Rugose corals of the Family Phillipsastreidae are abundant, diverse, and geographically widespread in the Frasnian (lower Upper Devonian) of western Canada. Species of the solitary genus Macgeea are described. Thamnophyllum and Peneckiella are branching forms. Biostratigraphic distribution of these species is reviewed, together with that of previously described Canadian massive phillipsastreid species belonging to the genera Phillipsastrea, Chuanbeiphyllum, Pachyphyllum, Smithicyathus, and Frechastraea. The coral biostratigraphy is expressed in terms of the Montagne Noire conodont zonation and modified western Canada rugose coral faunal assemblages.

Les coraux rugueux de la famille des Phillipsastreidae sont abondants, diversifiés et répandus sur le plan géographique dans le Frasnien (étage de base du Dévonien supérieur) de l'Ouest du Canada. L'ouvrage expose les espèces appartenant au genre solitaire Macgeea. Les genres Thamnophyllum et Peneckiella sont des formes branchues. La répartition biostratigraphique de ces espèces est examinée, ainsi que celle d'espèces massives de Phillipsastreidae canadiennes déjà décrites appartenant aux genres Phillipsastrea, Chuanbeiphyllum, Pachyphyllum, Smithicyathus et Frechastraea. La biostratigraphie des coraux est présentée suivant la zonation des conodontes de la Montagne Noire et les assemblages fauniques modifiés de coraux rugueux de l'Ouest du Canada.

To order books / Commandez des livres à :

Order Office / Bureau des commandes
NRC Research Press / Presses scientifique du CNRC National Research Council
Conseil national de recherches
Ottawa, ON K1A 0R6, Canada

Telephone / Téléphone : 613-990-2254
Fax / Télécopieur : 613-952-7656
E-mail / Courriel : pubs@nrc-cnrc.gc.ca
Order online / Commandez en ligne :
http://pubs.nrc-cnrc.gc.ca

Shipping and handling extra.
Canadian customers add GST (6%). Prepayment required unless order is accompanied by a purchase order number. Cheques or money orders should be made payable to: "Receiver General for Canada, Credit NRC". For all orders outside Canada, rates are in U.S. dollars. All sales are final.

Frais d'expédition et de manutention en plus.
Clients du Canada, ajoutez la TPS (6 %). Les commandes doivent être payées d'avance à moins qu'elles ne soient accompagnées d'un numéro de bon de commande. Les chèques ou mandats doivent être libellés au « Receveur général du Canada, crédit du CNRC ». Pour les commandes hors-Canada, les tarifs sont en dollars américains. Toutes les ventes sont finales.

NRC·CNRC

Canada

National Research Council Canada | Conseil national de recherches Canada

Flower Guide for Holiday Weekends

in Eastern Canada and Northeastern U.S.A.

By / Par
E.W. Larsen and B.I. Roots

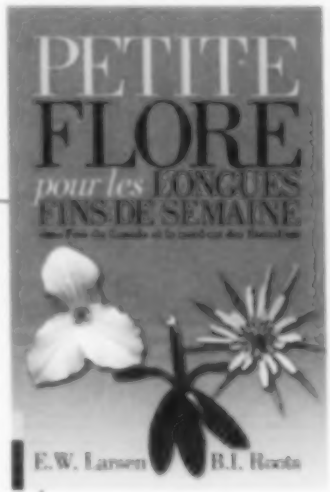
149 Pages 2005
Softcover / Édition brochée

ISBN 0-660-19096-6; NRC 46327
ISBN 0-660-19096-6; CNRC 46327

$14.95 CAN; Other countries: $14.95 US
14,95 $CAN; À l'étranger : 14,95 $US

Petite flore pour les longues fins de semaine

dans l'est du Canada et le nord-est des États-Unis



Par / By
E.W. Larsen et B.I. Roots

149 Pages 2005
Édition brochée / Softcover

ISBN 0-660-19096-6; CNRC 46327
ISBN 0-660-19096-6; NRC 46327

14,95 $CAN; À l'étranger : 14,95 $US
$14.95 CAN; Other countries: $14.95 US

To order books / Commandez des livres à :

Order Office / Bureau des commandes
NRC Research Press / Presses scientifique
du CNRC National Research Council
Conseil national de recherches
Ottawa, ON K1A 0R6, Canada

Telephone / Téléphone : 613-990-2254
Fax / Télécopieur : 613-952-7656
E-mail / Courriel : pubs@nrc-cnrc.gc.ca
Order online / Commandez en ligne :
http://pubs.nrc-cnrc.gc.ca

Shipping and handling extra.
Canadian customers add GST (6%). Prepayment required unless order is accompanied by a purchase order number. Cheques or money orders should be made payable to: "Receiver General for Canada, Credit NRC". For all orders outside Canada, rates are in U.S. dollars. All sales are final.

Frais d'expédition et de manutention en plus.
Clients du Canada, ajoutez la TPS (6 %). Les commandes doivent être payées d'avance à moins qu'elles ne soient accompagnées d'un numéro de bon de commande. Les chèques ou mandats doivent être libellés au « Receveur général du Canada, crédit du CNRC ». Pour les commandes hors-Canada, les tarifs sont en dollars américains. Toutes les ventes sont finales.

NRC·CNRC

Canada

National Research Council Canada
Ottawa, Canada K1A 0R6

Conseil national de recherches Canada
Ottawa, Canada K1A 0R6

Postage paid at Ottawa
Publications mail
Registration No. 40062591

Port payé à Ottawa
Poste-publication
Enregistrement n° 40062591

USPS periodical postage paid at Plattsburgh, NY 12901, USA

Canadian Journal of Fisheries and Aquatic Sciences

Volume 64, Number 8, August 2007

Journal canadien des sciences halieutiques et aquatiques

Volume 64, numéro 8, août 2007

	ARTICLES	ARTICLES
Brian K. Ginn, Brian F. Cumming, and John P. Smol	**1043–1054**	Assessing pH changes since pre-industrial times in 51 low-alkalinity lakes in Nova Scotia, Canada
Anders Nielsen and John R. Sibert	**1055–1068**	State–space model for light-based tracking of marine animals
Julie Deschênes, Marco A. Rodríguez, and Pierre Bérubé	**1069–1079**	Context-dependent responses of juvenile Atlantic salmon (*Salmo salar*) to forestry activities at multiple spatial scales within a river basin
Jennifer E. Roth, Kyra L. Mills, and William J. Sydeman	**1080–1090**	Chinook salmon (*Oncorhynchus tshawytscha*) — seabird covariation off central California and possible forecasting applications
Ana L. Ibañez, Ian G. Cowx, and Paul O'Higgins	**1091–1100**	Geometric morphometric analysis of fish scales for identifying genera, species, and local populations within the Mugilidae
Josh Korman, Caroline C. Melville, and Paul S. Higgins	**1101–1115**	Integrating multiple sources of data on migratory timing and catchability to estimate escapement for steelhead trout (*Oncorhynchus mykiss*)
Daniel C. Dauwalter, Dale K. Splinter, William L. Fisher, and Richard A. Marston	**1116–1129**	Geomorphology and stream habitat relationships with smallmouth bass (*Micropterus dolomieu*) abundance at multiple spatial scales in eastern Oklahoma
Sondre Aanes, Steinar Engen, Bernt-Erik Sæther, and Ronny Aanes	**1130–1142**	Estimation of the parameters of fish stock dynamics from catch-at-age data and indices of abundance: can natural and fishing mortality be separated?
Joseph H. Anderson and Thomas P. Quinn	**1143–1154**	Movements of adult coho salmon (*Oncorhynchus kisutch*) during colonization of newly accessible habitat
	ERRATUM	ERRATUM
Jason M. Cope and André E. Punt	**1155**	Erratum: Admitting ageing error when fitting growth curves: an example using the von Bertalanffy growth function with random effects

Front cover: Iboko Saito/Toshiyuki Tajima/Dex Image/Getty Images.

Page couverture : Iboko Saito/Toshiyuki Tajima/Dex Image/Getty Images.

Canada

Printed in Canada by / Imprimé au Canada par
NRC Reprographic Services / Service de reprographie CNRC
Page production by / Mise en page par
NRC Research Press / Presses scientifiques du CNRC

0706-652X(200708)64:8;1-K